Association studies and direct DNA sequencing implicate some known genetic susceptibility loci in the etiology of nonsyndromic orofacial clefts in sub-Saharan African populations

Gowans, Jephthah Joojo; Adeyemo, Wasiu L.; Eshete, Mekonen A.; Mossey, Peter; Busch, Tamara; Aregbesola, Babatunde

Published in:
Journal of Dental Research

DOI:
10.1177/0022034516657003

Publication date:
2016

Document Version
Peer reviewed version

Journal of
Dental Research

Association studies and direct DNA sequencing implicate some known genetic susceptibility loci in the etiology of nonsyndromic orofacial clefts in sub-Saharan African populations

Journal:	*Journal of Dental Research*
Manuscript ID	JDR-16-0113.R3
Manuscript Type:	Research Reports
Date Submitted by the Author:	n/a
Complete List of Authors:	Gowans, Lord; Kwame Nkrumah University of Science and Technology Mossey, P.A.; University of Dundee, Dental School Adeyemo, Wasiu; University of Lagos, Oral and Maxillofacial Surgery; Eshete, Mekonen; Addis Ababa University School of Medicine, Plastic Surgery Busch, Tamara; University of Iowa Carver College of Medicine, Craniofacial Anomalies Research Center Aregbesola, Babatunde ; Obafemi Awolowo University Teaching Hospital Complex Donkor, Peter; Kwame Nkrumah University of Science and Technology School of Medical Sciences Arthur, Fareed; Kwame Nkrumah University of Science and Technology College of Science Bello, Seidu ; State House Hospital Martinez, Andres; University of Iowa Roy J and Lucille A Carver College of Medicine Li, Mary; University of Iowa Roy J and Lucille A Carver College of Medicine Augustine-Akpan, Eno-Abasi ; University of Iowa Roy J and Lucille A Carver College of Medicine Daressa, Wakgari; Addis Ababa University School of Medicine Twumasi, Peter; Kwame Nkrumah University of Science and Technology College of Science Olutayo , James; University of Lagos M, Deribew ; Addis Ababa University School of Medicine Agbenorku, Pius; Kwame Nkrumah University of Science and Technology School of Medical Sciences Oti, Alex; Komfo Anokye Teaching Hospital, Oral and Maxillofcial Surgery Braimah, Rahman ; Obafemi Awolowo University Teaching Hospital Complex Plange-Rhule, Gyikua; Kwame Nkrumah University of Science and Technology School of Medical Sciences M, Gesses ; Yekatit 12 Hospital Obiri-Yeboah, Solomon; Kwame Nkrumah University of Science and Technology School of Medical Sciences Oseni, Ganiyu; LAUTECH Teaching Hospital

	Olaitan, Peter; LAUTECH Teaching Hospital Abdur-Rahman, Lukman ; University of Ilorin Teaching Hospital F, Abate; Yekatit 12 Hospital T, Hailu; Yekatit 12 Hospital P, Gravem; Haukeland Universitetssjukehus Ogunlewe, Mobolanle ; LAUTECH Teaching Hospital Buxo, Carmen; Universidad de Puerto Rico Recinto de Ciencias Medicas Marazita, Mary; University of Pittsburgh School of Medicine Adeyemo, Adebowale; Center for Research on Genomics and Global Health, National Human Genome Research Institute Murray, Jeffrey; The University of Iowa, Pediatrics Butali, Azeez; University of Iowa, Department of Oral Pathology, Radiology and Medicine
Keywords:	Craniofacial anomalies, Craniofacial biology/genetics, Genetics, Genomics, Orofacial cleft(s), Population genetics
Abstract:	Orofacial clefts (OFCs) are congenital dysmorphologies of the human face and oral cavity, with a global incidence of 1 per 700 live births. These anomalies exhibit multifactorial pattern of inheritance, with both genetic and environmental factors playing crucial roles. Many loci have been implicated in the aetiology of nonsyndromic cleft lip with or without cleft palate (NSCL/P) in populations of Asian and European ancestries through genome-wide association studies (GWAS) and candidate gene studies. However, few populations of African descent have been studied to date. Here, we show evidence of association of some loci with NSCL/P and nonsyndromic cleft palate only (NSCPO) in cohorts from Africa (Ghana, Ethiopia and Nigeria). We genotyped 48 SNPs that were selected from previous GWAS and candidate gene studies. These markers were successfully genotyped on 701 NSCL/P and 163 NSCPO cases, 1070 unaffected relatives and 1078 unrelated controls. We also directly sequenced 7 genes in 184 nonsyndromic OFC (NSOFC) cases and 96 controls from Ghana. Population-specific associations were observed in the case-control analyses of the sub-populations, with West African subpopulations (Ghana and Nigeria) showing similar pattern of associations. In meta-analyses of the case-control cohort, PAX7 (rs742071, p=5.10×10-03), 8q24 (rs987525, p=1.22×10-03) and VAX1 (rs7078160, p=0.04) were nominally associated with NSCL/P; MSX1 (rs115200552, p=0.01), TULP4 (rs651333, p=0.04), CRISPLD2 (rs4783099, p=0.02) and NOG1 (rs17760296, p=0.04) were nominally associated with NSCPO. Moreover, 7 loci exhibited evidence of threshold over-transmission in NSOFC cases in both transmission disequilibrium test (TDT) and family-based association for disease traits (DFAM) analyses. Through DNA sequencing, we also identified two novel, rare, potentially pathogenic variants (p.Asn323Asp and p.Lys426IlefsTer6) in ARHGAP29. In conclusion, we have shown evidence of association of many loci with NSCL/P and NSCPO. To the best of our knowledge, our study is the first to demonstrate any of these association signals in any African population.

Association studies and direct DNA sequencing implicate some known genetic susceptibility loci in the etiology of nonsyndromic orofacial clefts in sub-Saharan African populations

Lord Jephthah Joojo Gowans*[1,7,16,17], Wasiu L Adeyemo*[2], Mekonen Eshete*[3], Peter A Mosey[4], Tamara Busch[17], Babatunde Aregbesola[5], Peter Donkor[6,7], Fareed K N Arthur[1], Seidu A Bello[8], Andres Martinez[17], Mary Li[17], Eno-Abasi Augustine-Akpan[17], Wakgari Deressa[3], Peter Twumasi[1], Olutayo James[2], Deribew M[3], Pius Agbenorku[6,7], Alex Acheampong Oti[6,7], Rahman Braimah[5], Gyikua Plange-Rhule[7], Gesses M[9], Solomon Obiri-Yeboah[6,7], Ganiyu O Oseni[10], Peter B Olaitan[10], Lukman Abdur-Rahman[11], Abate F[9], Hailu T[9], Gravem P[12], Mobolanle Olugbenga Ogunlewe[10], Carmen J Buxó[13], Mary L Marazita[14], Adebowale A Adeyemo[15], Jeffrey C Murray[16], Azeez Butali[17]

[1]Department of Biochemistry and Biotechnology, Kwame Nkrumah University of Science and Technology (KNUST), Ghana, [2]College of Medicine, University of Lagos (CMUL), Nigeria, [3]Addis Ababa University, Ethiopia, [4]Department of Orthodontics, University of Dundee, Scotland, [5]Obafemi Awolowo University Teaching Hospital, Ile-Ife, Nigeria, [6]Department of Surgery, School of Medical Sciences, KNUST, Ghana, [7]Cleft Clinic, Komfo Anokye Teaching Hospital (KATH), Kumasi, Ghana, [8]Department of Oral and Maxillofacial Surgery, State House Hospital, Abuja, Nigeria, [9]Yekatit 12 Hospital medical college, Addis ababa ,Ethiopia, [10]Department of Burns and Plastic Surgery, Ladoke Akintola University of Technology Teaching Hospital, Osogbo, Nigeria, [11]Department of Pediatrics, University of Ilorin, Nigeria, [12]Haukeland University Hospital Bergen, Norway, [13]School of Dental Medicine, University of Puerto Rico Medical Sciences Campus, Puerto Rico, [14]Department of Oral Biology, University of Pittsburgh, USA, [15]Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA, [16]Department of Pediatrics,

*University of Iowa, USA, [17]Department of Oral Pathology, Radiology and Medicine, University of Iowa, USA. *Authors contributed equally to sample collection for this work*.

Corresponding Author: Lord Jephthah Joojo Gowans, Department of Biochemistry and Biotechnology, Kwame Nkrumah University of Science and Technology, PMB, University Post Office, Kumasi, Ghana, Africa. Tel.: +233 0244866389. Email: ljj.gowans@gmail.com.

Abstract

Orofacial clefts (OFCs) are congenital dysmorphologies of the human face and oral cavity, with a global incidence of 1 per 700 live births. These anomalies exhibit multifactorial pattern of inheritance, with both genetic and environmental factors playing crucial roles. Many loci have been implicated in the aetiology of nonsyndromic cleft lip with or without cleft palate (NSCL/P) in populations of Asian and European ancestries through genome-wide association studies (GWAS) and candidate gene studies. However, few populations of African descent have been studied to date. Here, we show evidence of association of some loci with NSCL/P and nonsyndromic cleft palate only (NSCPO) in cohorts from Africa (Ghana, Ethiopia and Nigeria). We genotyped 48 SNPs that were selected from previous GWAS and candidate gene studies. These markers were successfully genotyped on 701 NSCL/P and 163 NSCPO cases, 1070 unaffected relatives and 1078 unrelated controls. We also directly sequenced 7 genes in 184 nonsyndromic OFC (NSOFC) cases and 96 controls from Ghana. Population-specific associations were observed in the case-control analyses of the sub-populations, with West African subpopulations (Ghana and Nigeria) showing similar pattern of associations. In meta-analyses of the case-control cohort, *PAX7* (rs742071, p=5.10×10^{-03}), 8q24 (rs987525, p=1.22×10^{-03}) and *VAX1* (rs7078160, p=0.04) were nominally associated with NSCL/P; *MSX1* (rs115200552, p=0.01), *TULP4* (rs651333, p=0.04), *CRISPLD2* (rs4783099, p=0.02) and *NOG1* (rs17760296, p=0.04) were nominally associated with NSCPO. Moreover, 7 loci exhibited

evidence of threshold over-transmission in NSOFC cases in both transmission disequilibrium test (TDT) and family-based association for disease traits (DFAM) analyses. Through DNA sequencing, we also identified two novel, rare, potentially pathogenic variants (p.Asn323Asp and p.Lys426IlefsTer6) in *ARHGAP29*. In conclusion, we have shown evidence of association of many loci with NSCL/P and NSCPO. To the best of our knowledge, our study is the first to demonstrate any of these association signals in any African population.

Keywords

Africans, orofacial clefts, genetic heterogeneity, rare variants, Genome-Wide Association Studies (GWAS), candidate genes

Introduction

Human orofacial clefts (OFCs) are congenital malformations of the face and oral cavity, due to dysregulation of embryological processes. The global incidence of OFCs is 1 per 700 live births. However, race, ethnicity, geographical locations, environmental factors and socio-economic status influence the incidence of OFCs (Gorlin et al. 2001). The highest incidence occurs in Asians, followed by populations of European ancestry, whereas African populations have the lowest incidence (Mossey and Modell, 2012). Though there is no national prevalence data for Ghana and Ethiopia, a prevalence estimate of 0.5 per 1000 has been observed for Nigeria (Butali et al. 2014a). These observations presuppose that the relative contributions of individual susceptibility genes may vary across different human populations. OFCs may be syndromic or nonsyndromic, with the syndromic forms presenting with other congenital anomalies. The aetiology of the more common nonsyndromic OFCs (NSOFCs) is complex, exhibiting multifactorial pattern of inheritance. NSOFCs are classified into nonsyndromic cleft lip with or without cleft palate (NSCL/P) and nonsyndromic cleft palate only (NSCPO), and these

two groups have heterogeneous genetic architecture. NSCL/P comprises nonsyndromic cleft lip only (NSCL) and nonsyndromic cleft lip and palate (NSCLP) (Dixon et al. 2011).

To date, six genome-wide association studies (GWAS) and a meta-analysis have been published for NSOFCs, with these signals demonstrating association with NSCL/P but not NSCPO. In a GWAS involving Europeans, association was observed between a locus in Chr8q.24 and NSCL/P (Birnbaum et al. 2009). The 8q.24 signal was subsequently replicated in another GWAS of NSCL/P in Europeans from US (Grant et al. 2009). A third GWAS that involved cohorts of European ancestries also revealed that two additional loci, 17q22 (*NOG-1*) and 10q25 (*VAX1*), were associated with NSCL/P. Other loci yielded suggestive association with NSCL/P: 15q13.3 (*GREM1*), 13q31.1 (*SPRY2*) and 2p21 (*THADA*) (Mangold et al. 2010). Employing trios of Asian and European ancestries, a GWAS implicated 20q12 (*MAFB*) and 1p22.1 (*ABCA4*) in the aetiology of NSCL/P, with 17p13 (*NTN-1*) showing a suggestive association. Stratified analyses based on ancestries by the same GWAS showed that some signals were ancestry-specific: trios of European ancestry gave the strongest association for 8q.24 whereas those of Asian ancestry were strongly associated with *MAFB, ABCA4* and *IRF6* (Beaty et al. 2010). A meta-analysis also revealed additional NSCL/P susceptibility loci: *THADA, SPRY2,* 15q22.2 (*TPM1*) and 1p36 (*PAX7*) (Ludwig et al. 2012). Recently, a GWAS involving Asians implicated 16p13.3 (*ADCY9*) (Sun et al. 2015) in the aetiology of NSCL/P, whereas another GWAS involving dogs and Guatemala population gave a suggestive association for *ADAMTS20* (Wolf et al. 2015).

In both pre- and post-GWAS era, candidate gene and replication studies have been instrumental in identifying cleft susceptibility loci. Pathogenic variants in *IRF6* were shown to cause Van der Woude Syndrome (VWS) and Popliteal Pterygium Syndrome [PPS] (Kondo et al. 2002). Subsequently, a missense variant in *IRF6* (rs2235371) demonstrated over-transmission in NSCL/P cases of European ancestry (Zucchero et al. 2004). Another *IRF6* locus, rs642961, has also been shown to be associated with NSCL/P but not NSCPO (Rahimov et al. 2008).

Corollary to these observations, some studies (Kerameddin et al. 2015; Birnbaum et al. 2009) have confirmed a role of *IRF6* as NSCL/P risk loci in populations of Asian and European ancestries. Other candidate genes implicated in the aetiology of NSCL/P included *MSX1* (Rafighdoost et al. 2013), *BMP4* (Suzuki et al. 2009), *FOXE1* (Moreno et al. 2009), *AXIN2* (Letra et al. 2012), *CRISPLD2* (Chiquet et al. 2007), *NOG1* and *FGFR2* (Leslie et al. 2015).

Among Africans, genetic studies on OFCs are limited. A study involving a Nigerian cohort implicated *MSX1*, but not other loci, in the aetiology of NSCL/P (Butali et al. 2011). Other studies that recruited Kenyans (Wheatherley-White et al. 2011) and Congolese (Figueiredo et al. 2014) could not replicate the association for cleft susceptibility loci among Africans, probably due to small sample size and population heterogeneity. Moreover, sequencing of GWAS loci in cohorts from Ethiopia and Nigeria reported some rare, potentially causative variants (Butali et al. 2014b). Conducting genetic and genomics studies using cleft cohort from Africa may identify novel and population-specific signals. However, it is also important for us to investigate the role of identified signals and biologically relevant genes from existing European and Asian studies in the African population. The present study was aimed at replicating the association between reported GWAS and candidate gene loci in our NSCL/P cohort. We also tested the hypothesis that NSCL/P loci may also contribute to NSCPO susceptibility in Africans. Finally we screened for rare, potentially pathogenic variants in 7 candidate genes at risk loci that are usually associated with NSCL/P.

Subjects and Methods

We recruited 3,585 participants from Ghana, Ethiopia and Nigeria (Table 1; Supplemental Methods). All sample and data collection at various study sites were approved by the local Institutional Review Boards: KATH (Ghana) – CHRPE/AP/217/13, CMUL (Nigeria) – ADM/DCST/HREC/APP/1374 and Addis Ababa University Teaching Hospital (Ethiopia) -

3.10/027/2015. Before sample and data collection, written, informed consent was obtained from each participating family. DNA processing is shown in Supplemental Methods.

SNP Selection

We selected SNPs with MAF of 5% and above in the African population for genotyping; these have either been previously reported in peer review journals or were identified in animal studies and during our re-sequencing studies. These include SNPs that are associated with NSCL/P in candidate genes studies and GWAS in European and Asian populations (Supplemental Table S1).

SNP Genotyping

We genotyped 48 SNPs (Supplemental Table S1) on a total of 3,585 samples - 872 NSOFC cases (163 NSCPO, 340 NSCL, 361 NSCLP, and 8 “un-typed”), 1635 unaffected relatives and 1078 unrelated controls, using 192.24 Fluidigm SNP Genotyping Protocol (Supplemental Methods). The “un-typed” (samples from probands) and other samples, however, failed quality control checks and were not included in the final statistical analyses (Table 1).

Statistical Analyses for association studies

During quality control checks, we resolved Mendelian errors in case-parent triads and dropped from the final analyses samples that were not successfully genotyped on at least 95% of the 48 genotyped SNPs. We computed Hardy Weinberg Equilibrium (HWE) using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/). We then conducted case-control analyses to determine association in each subpopulation and meta-analyses of the three subpopulations based on Table 1. For this test, we used $p<0.05$ to denote nominal association and a Bonferroni Correction of 141 tests to ascertain a threshold for formal significance of $p=3.54\times10^{-4}$. The 141 tests comprised of 47 SNPs that passed HWE × 3 cleft sub-phenotypes × 1 racial group x 1

test. Out of the 48 SNPs, only one failed HWE (p<0.05). Additional analyses to determine over-transmission of the rare alleles were conducted using the Transmission Disequilibrium Test (TDT) and Family-Based Association for Disease Traits (DFAM). TDT used only the case-parent triad information (Table 1) while DFAM allowed us to combine both triad and dyad data. For these tests, the significant p-value was 0.05. Parent of Origin (POO) effects, and gene-gene interactions (epistasis) was also calculated. The probands in the case-control arm of the study (Table 1) are the same probands in the family-based studies.

DNA Sequencing

We directly sequenced *VAX1*, *PAX7*, *ARHGAP29*, *MSX1*, *FOXE1*, *BMP4* and *MAFB* in 184 NSOFC cases (131 NSCL/P and 53 NSCPO) from Ghana using Sanger Sequencing (Supplemental Methods; Butali et al. 2014b). We also performed segregation analyses on observed potentially pathogenic missense, frameshift and splice site variants by sequencing available parental samples. We further sequenced 96 unrelated Ghanaian controls to ascertain whether the novel variants we encountered in NSOFC cases also occurred in controls or not.

Results

Association Analyses

In meta-analyses of the case-control cohorts from the three subpopulations, we successfully demonstrated nominal association between *PAX7* (rs742071, p=5.10×10^{-03}), 8q24 (rs987525, p=1.22×10^{-03}) as well as *VAX1* (rs7078160, p=0.04) and NSCL/P; *MSX1* (rs115200552, p=0.01), *TULP4* (rs651333, p=0.04), *CRISPLD2* (rs4783099, p=0.02) and *NOG1* (rs17760296, p=0.04) were nominally associated with NSCPO (Table 2), with the direction of effect being the same as reported by earlier studies. Among Ethiopians (Supplemental Table S2), *PAX7* (rs742071, p=5.57×10^{-03}), *IRF6* (rs642961, p=0.02), *DYSF* (rs2303596, p=2.31×10^{-03}), 8q24 (rs987525, p=7.82×10^{-04}) and *MAFB* (rs13041247 and rs11696257, all with p=0.04)

were nominally associated with NSCL/P; *ABCA4* (rs481931 and rs4147811, all with p=0.03) and *NTN1* (rs8081823, p=0.03) were nominally associated with NSCPO. Moreover, subphenotype analyses of the Ethiopian NSCL/P cohort showed that the *PAX7*, *DYSF*, *MSX1*, *SPRY2* (rs9574565, $p=7.05\times10^{-03}$) and *MAFB* signals were particularly stronger for NSCL whereas the *IRF6* (rs642961, $p=9.11\times10^{-03}$) and 8q24 (rs987525, $p=1.07\times10^{-03}$) signals were stronger for NSCLP (Supplemental Table S2). Among Ghanaians (Supplemental Table S3), *ABCA4* (rs560426, p=0.03) and *VAX1* (rs7078160, p=0.03) were nominally associated with NSCLP, with subphenotype analyses of the NSCL/P cohort showing that the *ABCA4* locus was strongly associated with NSCLP. *ABCA4* (rs4147811, $p=7.48\times10^{-03}$) and *CRISPLD2* (rs4783099, p=0.04) were nominally associated with NSCL/P and NSCPO, respectively, among Nigerians (Supplemental Table S4). Subphenotype analyses of the Nigerian NSCL/P (Supplemental Table S4) showed that *PAX7* (rs742071, p=0.02) and *ARHGAP29* (rs138751793, p=0.04) signals were stronger for NSCL whereas another SNP at the *ABCA4* locus (rs481931, $p=2.87\times10^{-03}$) was strongly associated with NSCLP. However, none of these case-control associations passed Bonferroni correction.

For TDT and DFAM (Tables 3 and 4) for all the three subpopulations, seven loci demonstrated formal significance with NSOFCs at $p\leq0.05$. Formal significance for TDT and DFAM was evaluated at $p\leq0.05$ because these are secondary analyses compared with case-control analyses, and are not true independent tests. All family-based studies suggested that the minor allele of *ABCA4* (rs560426) was over-transmitted in NSCLP cases among Africans. *PAX7* (rs742071) also consistently showed evidence of over-transmission in NSCL cases in both TDT and DFAM. *MSX1* (rs115200552) and *AXIN2* (rs3923086) also demonstrated strong over-transmission in NSCLP cases in DFAM analyses whereas *MTHFR* (rs1801131) and *DYSF* exhibited over-transmission in NSCL cases in TDT and DFAM analyses, respectively. Only a SNP of *VAX1* demonstrated over-transmission in NSCPO cases.

Parent of Origin Effects

Parent-of-origin (POO) effects were not observed for almost all SNPs, except rs16260 of *CDH1*. For rs16260, a trend towards association (p=0.0764) was observed for all clefts. The rs16260 SNP exhibited a maternal imprinting or maternal over-transmission effect.

Gene-Gene Interactions

In Gene-Gene (G×G) or epistatic interactions, three SNPs exhibited evidence of epistasis with other SNPs. Each of these epistatic interactions yielded p=0.02. A SNP for *ABCA4*, rs560426, interacted with Chr6 rs2674394 (gene desert). Moreover, rs2303596 of *DYSF* interacted with rs3923086 of *AXIN2*. Finally, rs8069536 of *NTN1* interacted with rs17820943, rs13041247 and rs11696257, all of *MAFB*. However, none of these G×G interactions passed Bonferroni correction.

Direct DNA sequencing of seven selected genes

We observed several rare and/or novel variants in the 7 genes that we sequenced (Table 5, Supplemental Table S5). Rare variants, as used here, refer to either a novel variant or a variant whose MAF is less than or equals to 1%. Some of these variants were predicted to be potentially pathogenic by various bioinformatics tools whereas others were depicted as benign. A *de novo* occurrence could not be demonstrated for any of these variants because either the variant was present in at least one parent or not both parents were available for segregation analysis. Lastly, some of the novel variants we observed occurred in controls (e.g. all *VAX1* variants) whereas others were not observed in controls (e.g. all *ARHGAP29* variants).

Discussion

We have successfully demonstrated associations (both nominal in case-control analyses and threshold in TDT and DFAM analyses) between some loci and NSCL/P in cohorts from

Africa. We also tested the hypothesis that these loci also contribute to NSCPO in Africans and observed some interesting associations. The 8q24 locus exhibited the strongest nominal significance with NSCL/P in case-control meta-analyses, with the trends suggesting this locus may be relevant in all three subpopulations. The test of heterogeneity also suggested largely the absence of heterogeneity at this locus among the three African populations. We observed that among Africans, the associated minor C allele of rs987525 (http://browser.1000genomes.org) conferred reduced susceptibility while the major A allele is the risk allele. Irrespective of these differences in minor alleles, our result is in harmony with earlier studies (Birnbaum et al. 2009; Grant et al. 2009; Mangold et al. 2010; Beaty et al. 2010; Ludwig et al. 2012) that demonstrated that the A allele of rs987525 is a risk allele for NSCL/P in Europeans. These observations suggest that the actual risk variant(s) is/are in linkage disequilibrium (LD) with the A allele of rs987525. Fine mapping of the African Haplotype (which is smaller in the 8q24 region) will help identify the risk variant(s). Our observations corroborate those made elsewhere (Beaty et al., 2010; Murray et al., 2012) that suggested that the varied ethnic association of the rs987525 allele largely depends on its MAF in various populations. Current evidence suggests that though the 8q24 window is a gene desert, it harbors very remote *cis*-acting craniofacial enhancer elements that regulate the expression of oncogenic *MYC* in the developing face; perturbation of this regulatory network leads to craniofacial dysmorphologies, including sporadic CL/P, in mice (Uslu et al. 2014).

The C677T (rs1801133) SNP of *MTHFR*, but not A1298C (rs1801131), has largely been associated with reduced risk for NSCL/P in Asians (Martinelli et al. 2015; Pan et al. 2015; Zhao et al. 2014) and to some extent, in European-derived populations (Estandia-Ortega et al. 2014; de Aguiar et al. 2015), though not all studies (Sozen et al. 2009) replicated the association. Interestingly, we have demonstrated in TDT analyses that *MTHFR* is significantly associated with NSCL among Africans and that it is the C minor allele of A1298C (rs1801131) SNP that confers a reduced risk, suggesting A is the risk allele. *AXIN2* has been implicated in the

aetiology of NSOFCs in multiple populations, except Africans, with rs3923086 demonstrating association with NSCLP among Asians (Letra et al. 2012). Other studies (Mostowska et al. 2012; Araujo et al. 2015) have replicated the association between *AXIN2* and NSCL/P. Here, we have demonstrated that rs3923086 (*AXIN2*) is also associated with NSCLP among Africans in DFAM analyses. Other candidate genes (e.g. *DYSF*) also showed evidence of association with NSOFCs among Africans, buttressing the relevance of this approach in aetiologic "gene hunting".

Other SNPs, other than already reported ones, may be responsible for the reported associations between certain loci and NSOFCs in some ethnicities. Through direct DNA sequencing of *MSX1* gene, we observed over-transmission of the minor allele of rs115200552 in NSOFC cases. Subsequent genotyping of this SNP in 3,585 individuals showed that this SNP was associated with NSCPO (p=0.01) in case-control meta-analyses, though family-based studies also suggest this marker may also be a risk allele for NSCLP. Earlier studies involving Africans from Nigeria implicated *MSX1* in the aetiology of NSCL/P (Butali et al. 2011).

We could not detect formal association between some GWAS and candidate gene loci and NSCL/P, presupposing either these loci may not play a role in the aetiology of NSCL/P in Africans or the genotyped SNPs may not be the tag SNPs for Africans. Lack of statistical power due to sample size and low MAF of the genotyped SNPs in Africans could also be possible reasons. For example, a SNP, rs2235371 of *IRF6* which is in high LD and same locus as rs642961, that has been associated with NSCL/P mostly among Asians (Sun et al. 2015) and in some Europeans (Zucchero et al. 2004), does not exist in the African population (http://browser.1000genomes.org/index.html). It is also possible that even when no associations are detected between reported loci and NSOFCs, potentially pathogenic variants may be observed in NSOFC cases. Therefore, GWAS and whole genome sequencing (WGS) of NSOFC cases from Africa is required to detect more risk loci.

Subphenotype and sub-population analyses (even among the same racial group) may be crucial in detecting association between certain loci and NSOFCs. In both TDT and DFAM analyses, we observed that rs560426 of *ABCA4* was associated with NSCLP but not the other OFC subphenotypes. Case-control analyses further suggested that the *ABCA4* locus may be crucial in NSOFC aetiology in all three African populations. *PAX7* (rs742071) exhibited nominal association with NSCL/P in case-control meta-analyses, with subpopulation analyses suggesting this signal originated mainly from the Ethiopian and Nigerian cohorts which exhibited some level of heterogeneity. However, TDT and DFAM subphenotype analyses demonstrated that rs742071 exhibited over-transmission in NSCL cases in all three populations. In case-control meta-analyses, *VAX1* (rs7078160) was nominally associated with NSCL/P, with subpopulation analyses suggesting the two West African countries (largely Ghana) drive this signal.

Rare variants, but not necessarily common variants, may account for the link between certain loci and NSOFCs. We observed many missense and a frameshift mutations in sequenced genes. No *de novo* occurrence was observed for any of these variants due to unavailability of some parental samples. Moreover, some of the novel variants were also observed in some clinically unaffected parents and controls. We sequenced the novel variants in 96 controls from Ghana and the likelihood of identifying these novel variants in more controls (i.e. >96) is possible. Nonetheless, these variants are absent in over 1000 individuals in 1000genomes database (with over 300 Africans), over 61,000 individuals in ExAC database as well as 6500 individuals in EVS. There is also the need to functionally validate the pathogenicity or otherwise of these variants *in vivo*. Rare variants in *ARHGAP29* (Leslie et al. 2012), *PAX7* and *VAX1* (Butali et al. 2013; Leslie et al. 2015), *BMP4* (Suzuki et al. 2009), *FOXE1* (Moreno et al. 2009), *MAFB* (Butali et al. 2014b) and *MSX1* (Liang et al. 2012) have been observed in NSOFC cases.

The incidence of OFC in Africans is much lower than in Europeans and Asians (Mossey and Modell, 2012; Butali et al. 2014a), even though these populations may share the same or similar genetic susceptibility loci for OFCs, as observed in the present study. Though under-ascertainment due to lack of birth defect registries in most African countries could contribute to the low incidence (Butali et al. 2014a), the low incidence of OFCs among Africans may be real, as African-derived populations in the Caribbean have lower OFC incidence that is similar to their ancestral population (Mossey and Modell, 2012). We therefore hypothesize the possible existence of genetic protective variants in the African genome, whose "rescue mission" reduces clefting. The identification and elucidation of such protective variants can be translated to European and Asian populations to bring about reduced OFC incidence, and eventually prevention.

Conclusion

The present study has shown evidence of association of certain loci with NSOFCs at both nominal and threshold significance. For instance, we have for the first time shown that the 8q.24 locus is a risk locus in Africans. Our study has thus corroborated earlier suggestion that the 8q24 locus may be a risk locus for NSCL/P across major ethnicities, though the effect size is smaller in Asians due to lower MAF. Subphenotype as well as sub-population analyses and genotyping of other SNPs, other than those already reported for some loci, may be crucial in identifying NSOFC loci in various ethnicities and populations. We have also demonstrated the existence of rare variants, both novel and known ones, in NSOFC cases from Africa. In conclusion, we have for the first time demonstrated associations between the SNPs we studied and NSOFC among Africans. Our study is crucial for understanding the genetic architecture of NSOFCs in Africans and further suggests the need to carry out GWAS and WGS for every ethnicity as far as complex traits are concerned.

Acknowledgements

We are grateful to of all families that participated in this research. We also want to appreciate all Doctors and Nurses at the various Teaching Hospitals where sample collection were carried out: Komfo Anokye Teaching Hospital (Ghana), University of Lagos Teaching Hospital (Nigeria), Obafemi Awolowo University Teaching Hospital (Nigeria), Ladoke Akintola University of Technology Teaching Hospital (Nigeria) and Yekatit 12 Hospital Medical college (Ethiopia). Contributions of members of Murray and Butali Laboratories, University of Iowa, are acknowledged. This work was supported by NIDCR R00 Grant DE022378 and Robert Wood Johnson Foundation Grant number 72429 (AB), NIH R37 grants DE-08559 and DE-016148 (JCM), The Ghana Cleft Foundation (LJJG), R01-DE009886 and R01-DE012472 (MLM) and U54-MD007587, R25-MD007607 and K99-DE024571 (CJB). Authors have no conflict of interests to declare.

References

Araujo TK, Secolin R, Felix TM, Souza LT, Fontes MI, Monlleo IL, Souza J, Fett-Conte AC, Secolin R, Lopes-Cendes I, et al. 2015. A multicentric association study between 39 genes and nonsyndromic cleft lip and palate in a Brazilian population. J Craniofacial Surg. S1010-5182(15)00252-00258.

Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, Hetmanski JB, Liang KY, Wu T, Murray T, Fallin MD, et al. 2010. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near *MAFB* and *ABCA4*. Nat Genet. 42(6):525–529.

Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, Rubini M, Baluardo C, Ferrian M, Almeidade Assis N, Alblas MA, et al. 2009. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. Nat Genet. 41(4):473–477.

Butali A, Mossey PA, Adeyemo WL, Jezewski PA, Onwuamah CK, Ogunlewe MO, Ugboko VI, Adejuyigbe O, Adigun AI, Abdur-Rahman LO, et al. 2011. Genetic studies in the Nigerian population implicate an *MSX1* mutation in complex oral facial clefting disorders. Cleft Palate Craniofac J. 48(6):646–653.

Butali A, Suzuki S, Cooper ME, Mansilla AM, Cuenco K, Leslie EJ, Suzuki Y, Niimi T, Yamamoto M, Ayanga G, et al. 2013. Replication of genome wide association identified candidate genes confirm the role of common and rare variants in *PAX7* and *VAX1* in the etiology of nonsyndromic CL[P]. Am J Med Genet A 161A(5):965–972.

Butali A, Adeyemo WL, Mossey PA, Olasoji HO, Onah II, Adebola A, Efunkoya, Akintububo A, James O, Adeosun OO, et al. 2014a. Prevalence of orofacial clefts in Nigeria. Cleft Plate Craniofac J. 51(3):320-325.

Butali A, Mossey P, Adeyemo WL, Eshete M, Gaines LAL, Braimah RO, Aregbesola BS, Rigdon J, Emeka C, Olutayo J, et al. 2014b. Rare Functional Variants in Genome–wide Association Identified Candidate Genes for Non-syndromic Clefts in the African Population. Am J Med Genet A 164(10):2567–2571.

Butali A, Mossey P, Tiffin N, Adeyemo W, Eshete M, Mumena C, Audu R, Onwuamah C, Agbenorku P, Ogunlewe M, et al. 2015. Multidisciplinary approach to genomics research in Africa: the AfriCRAN model. Pan Afr Med J. 21:229.

Chiquet BT, Lidral AC, Stal S, Mulliken JB, Moreno LM, Arcos-Burgos M, Valencia-Ramirez C, Blanto, SH, Hetch JT. 2007. *CRISPLD2*: a novel NSCLP candidate gene. Hum Mol Genet. 16(18):2241-2248.

de Aguiar PK, Coletta RD, de Oliveira AM, Machado RA, Furtado PG, de Oliveira LA, de Aquino SN, Martelli-Junior H, de Almeida Reis SR, Moreira HS, et al. 2015. Rs1801133C>T polymorphism is a risk factor for nonsyndromic cleft lip with or without cleft palate in the Brazilian population. Birth Defects Res A Clin Mol Teratol. 103(4):292-298.

Dixon MJ, Marazita ML, Beaty TH, Murray JC. 2011. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet. 12(3):167–178.

Estandia-Ortega B, Velazquez-Aragon JA, Alcanta-Ortigoza MA, Reyna-Fabian ME, Villagomez-Martinez S, Gonzalez-Del Angel A. 2014. 5,10-Methylenetetrahydrofolate reductase single nucleotide polymorphisms and gene-enevironment interaction analysis in non-syndromic cleft lip/palate. Eur J Oral Sci. 122(2):109-113.

Figueiredo JC, Ly S, Raimondi H, Magee K, Baurley JW, Sanchez-Lara PA, Ihenacho U, Yao C, Edlund CK, van den Berg D, et al. 2014. Genetic risk factors for orofacial clefts in Central Africans and Southeast Asians. Am J Med Genet A. 164A(10):2572-2580.

Gorlin RJ, Cohen MM, Hennekam RCM. 2001. Syndromes of the neck and head. Oxford University Press, New York.

Grant SFA, Wang K, Zhang H, Glaberson W, Annaiah K, Kim CE, Bradfield PJ, Glessner JT, Thomas KA, Garris M, et al. 2009. A genome-wide association study identifies a locus for non-syndromic cleft lip with or without cleft palate on 8q.24. J Pediatr. 155(6):909–913.

Kerameddin S, Namipashaki A, Ebrahimi S, Ansari-Pou N. 2015. *IRF6* Is a Marker of Severity in Nonsyndromic Cleft Lip/Palate. J Dent Res. 94(9 Suppl):226S-232S.

Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, Howard E, Ferreira de Lima RLL, Daack-Hirsch S, Sander A, et al. 2002. Mutations in *IRF6* cause van der Woude and popliteal pterygium syndromes. Nat Genet. 32(2):285–289.

Leslie EJ, Mansilla MA, Biggs LC, Schuette K, Bullard S, Cooper M, Dunnwald M, Lidral AC, Marazita ML, Beaty TH, et al. 2012. Expression and mutation analyses implicate *ARHGAP29* as the etiologic gene for the cleft lip with or without cleft palate locus identified by genome wide association on chromosome 1p22. Birth Defects Res A Clin Mol Teratol. 94(11):934–942.

Leslie EJ, Taub MA, Liu H, Steinberg KM, Koboldt DC, Zhang Q, Carlson JC, Hetmanski JB, Wang H, Larson DE, et al. 2015. Identification of Functional Variants for Cleft Lip with or without Cleft Palate in or near *PAX7, FGFR2*, and *NOG* by Targeted Sequencing of GWAS Loci. Am J Hum Genet. 96(3):397–411.

Letra A, Bjork B, Cooper ME, Szabo-Rogers H, Deleyiannis FW, Field LL, Czeizel AE, Ma L, Garlet GP, Poletta FA, et al. 2012. Association of *AXIN2* with non-syndromic oral clefts in multiple populations. J Dent Res. 91(5):473-478.

Liang J, Zhu L, Meng L, Chen D, Bian Z. 2012. Novel nonsense mutation in *MSX1* causes tooth agenesis with cleft lip in a Chinese family. Eur J Oral Sci 120(4):278–282.

Ludwig KU, Mangold E, Herms S, Nowak S, Reutter H, Paul A, Becker J, Herberz R, Alchawa T, Nasser E, et al. 2012. Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. Nat Genet. 44(9):968–971.

Mangold E, Ludwig KU, Birnbaum S, Baluardo C, Ferrian M, Herms S, Reutter H, de Assis NA, Chawa TA, Mattheisen M, et al. 2010. Genomewide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. Nat Genet. 42(1):24–26.

Martinelli M, Girardi A, Cura F, Nouri N, Pinto V, Carinci F, Morselli PG, Salehi M, Scapoli L. 2015. Non-syndromic cleft lip with or without cleft palate in Asian populations: Association analysis on three gene polymorphisms of the folate pathway. Arch Oral Biol. 61:79-82.

Moreno LM, Mansilla MA, Bullard SA, Cooper ME, Busch TD, Machida J, Johnson MK, Brauer D, Krahn K, Daack-Hirsch S, et al. 2009. *FOXE1* association with both isolated cleft lip with or without cleft palate, and isolated cleft palate. Hum Mol Genet. 18(24):4879–4896.

Mossey PA, Modell B. 2012. Epidemiology of oral clefts 2012: an international perspective. Front Oral Biol. 16:1–18.

Mostowska A, Hozyasz KK, Wojcicki P, Lasota A, Dunin-Wilczynskal I, Jagodzinski PP. 2012. Association of *DVL2* and *AXIN2* gene polymorphisms with cleft lip with or without cleft palate in a Polish population. Birth Defects Res A Clin Mol Teratol. 94(11):943-9450.

Murray T, Taub MA, Ruczinski I, Scott AF, Hetmanski JB, Schwender H, Patel P, Zhang TX, Munger RG, Wilcox AJ, et al. 2012. Examining markers in 8q24 to explain differences in evidence for association with cleft lip with/without cleftpalate between Asians and Europeans. Genet Epidemiol. 36(4): 392-399.

Pan X, Wang P, Yin X, Liu X, Li D, Li X, Wang Y, Li H, Yu Z. 2015. Association between Maternal *MTHFR* Polymorphisms and Nonsyndromic Cleft Lip with or without Cleft Palate in Offspring, A Meta-Analysis Based on 15 Case-Control Studies. 2015. Int J Fertil Steril. 8(4):463-480.

Pegelow M, Koillinen H, Magnusson M, Fransson I, Unneberg P, Kere J, Karsten A, Peyrard-Janvid M. 2014. Association and Mutation Analyses of the IRF6 Gene in Families With Nonsyndromic and Syndromic Cleft Lip and/or Cleft Palate. Cleft Palate Craniofac J.; 51(1):49–55.

Rafighdoost H, Hashemi M, Narouei A, Eskanadri-Nasab E, Dashti-Khadivaki G, Taheri M. 2013. Association Between *CDH1* and *MSX1* Gene Polymorphisms and the Risk of Nonsyndromic Cleft Lip and/or Cleft Palate in a Southeast Iranian Population. Cleft Palate Craniofac J. 50(5): e98–e104.

Rahimov F, Marazita ML, Visel A, Cooper ME, Hitchler MJ, Rubini M, Domann FE, Govil M, Christensen K, Bille C, et al. 2008. Disruption of an *AP-2α* binding site in an *IRF6* enhancer is associated with cleft lip. Nat Genet. 40(11):1341–1347.

Sozen MA, Tolarova MM, Spritz RA. 2009. The common *MTHFR* C677T and A1298C variants are not associated with the risk of non-syndromic cleft lip/palate in northern Venezuela. J Genet Genomics 36(5): 283-288.

Sun Y, Huang Y, Yin A, Pan Y, Wang Y, Wang C, Du Y, Wang M, Lan F, Hu Z, et al. 2015. Genome-wide association study identifies a new susceptibility locus for cleft lip with or without a cleft palate. Nat Commun. 6:6414.

Suzuki S, Marazita ML, Cooper ME, Miwa N, Hing A, Jugessur A, Natsume N, Shimozato K, Ohbayashi N, Suzuki Y, et al. 2009. Mutations in *BMP4* are associated with subepithelial, microform, and overt cleft lip. Am J Hum Genet. 84(3):406–411.

Uslu VV, Petretich M, Ruf S, Langenfeld K, Fonseca NA, Marioni JC, Spitz F. 2014. Long-range enhancers regulating Myc expression are required for normal facial morphogenesis. Nat Genet. 46(7):753-758.

Weatherley-White RC, Ben S, Jin Y, Riccardi S, Arnold TD, Spritz RA. 2011. Analysis of genome-wide association signals for nonsyndromic cleft lip/palate in a Kenya African Cohort. Am J Med Genet A. 155A(10):2422–2425.

Wolf ZT, Brand HA, Shaffer JR, Leslie EJ, Arzi B, Willet CE, Cox TC, McHenry T, Narayan N, Feingold E, et al. 2015. Genome-Wide Association Studies in Dogs and Humans Identify *ADAMTS20* as a Risk Variant for Cleft Lip and Palate. PLoS Genet 11(3): e1005059.

Zhao M, Ren Y, Shen L, Zhang Y, Zhou B. 2014. Association between *MTHFR* C677T and A1298C polymorphisms and NSCL/P risk in Asians: a meta-analysis. PLoS One 9(3):e88242.

Zucchero TM, Cooper ME, Maher BS, Daack-Hirsch S, Nepomuceno B, Ribeiro L, Caprau D, Christensen K, Suzuki Y, Machida J, et al. 2004. Interferon regulatory factor 6 (*IRF6*) gene variants and the risk of isolated cleft lip or palate. N Engl J Med. 351(8):769-780.

Table 1: Subphenotypes, gender and sample types of study cohort that passed quality control checks and were included in statistical analyses

	Number of samples per population			
Cleft Subphenotype of probands	Ghana	Ethiopia	Nigeria	**Total**
	Case control cohort			
NSCL	162	101	77	340
NSCLP	144	143	74	361
NSCPO	102	21	40	163
Unrelated Controls	408	357	313	1078
	Case-parents trios			
NSCL	52	2	20	74
NSCLP	48	3	26	77
NSCPO	34	1	7	42
	Case-parent dyads			
NSCL	77	84	51	212
NSCLP	76	134	47	257
NSCPO	53	20	32	105
	other trios			
NSCL	18	0	0	18
NSCLP	14	0	0	14
NSCPO	11	0	0	11
	other dyads			
NSCL	8	0	0	8
NSCLP	3	0	0	3
NSCPO	3	0	0	3
	Singletons			
NSCL	5	13	6	24
NSCLP	1	8	1	10
NSCPO	2	0	1	3
	Tetrads			
NSCLP	2	0	0	2
	Pentads			
NSCLP	1	0	0	1

Case probands consisted of 423 males and 441 females whereas unrelated controls were made up of 441 males and 637 females. The probands in the case-control arm of the study are the same probands in the family-based studies. In some of the designated singletons, parental samples failed data cleaning and were dropped from statistical analyses, hence the designation of such families as singletons. Singletons were informative in the case-control arm of our study but not the family-based studies. Tetrads and pentads were collected from families where two individuals were affected with clefts. “Other trios and dyads” largely refers to case-mother-maternal grandmother trios, case-mother-sibling trios as well as case-siblings trios and dyads. Case-parent trios, tetrads and pentads were employed in Transmission Disequilibrium Test (TDT) whereas all sample types, except singletons and unrelated controls, were used for Family-Based Association for Disease Traits (DFAM) analyses. Only case probands and unrelated controls were included in the case-control analyses.

Table 2: Meta-analyses of the case-control cohorts from Ghana, Ethiopia and Nigeria

Part A: Meta-analyses of NSCL/P and NSCPO case-control cohorts from all three countries									
SNP	Probable gene/loci	Minor alleles[a]	African MAF	NSCL/P			NSCPO		
				p	OR	I	*p*	OR	I
rs1801131	*MTHFR*	C/A[r]	0.15	0.32	1.08	0.00	0.19	0.79	0.00
rs1801133	*MTHFR*	A/G[p]	0.09	0.49	1.08	18.19	0.44	0.83	0.00
rs766325	*PAX7*	G/A[c,d,r]	0.18	0.29	0.92	0.00	0.23	0.82	0.00
rs742071	*PAX7*	T/G[r]	0.39	**5.10E-03[f]**	**1.19**	**54.68**	0.76	0.96	0.00
rs560426	*ABCA4*	C/T[b,r]	0.49	0.10	0.90	6.15	0.16	1.18	0.00
rs481931	*ABCA4*	T/G[p]	0.10	0.40	1.09	11.13	0.49	0.85	0.00
rs4147811	*ABCA4*	T/C[p]	0.11	0.23	1.13	67.35	0.93	1.02	0.00
rs138751793	*ARHGAP29*	C/T[e]	0.02	0.24	1.32	0.00	0.47	1.34	27.90
rs6677101	*SLC25A24*	G/T[b,d,r]	0.33	0.80	0.98	12.11	0.87	1.02	53.89
rs861020	*IRF6*	A/G[r]	0.11	0.23	1.11	0.00	0.83	0.96	24.15
rs34743335	*IRF6*	T/A	0.02	0.59	0.90	0.00	0.84	0.89	38.34
rs642961	*IRF6*	A/G[r]	0.09	0.32	1.11	68.47	0.57	0.88	44.17
rs7590268	*THADA*	G/T[r]	0.20	0.74	0.98	0.00	0.38	0.87	0.00
rs4332945	*DYSF*	T/G[b,d,r]	0.16	0.94	0.99	0.00	0.97	1.01	0.00
rs2303596	*DYSF*	T/C[c,d,p]	0.22	0.20	0.91	75.32	0.57	1.09	73.54
rs227782	*DYSF*	A/G[b,r]	0.42	0.33	1.06	0.00	0.35	1.12	61.90
rs115200552	*MSX1*	C/G[e]	0.02	0.38	1.16	28.63	**0.01[f]**	**1.81**	**0.00**
rs12532	*MSX1*	G/A[d,p]	0.44	0.49	0.96	0.00	0.37	0.90	0.43
rs2674394	Gene Desert	A/C[r]	0.17	0.62	1.04	0.00	0.68	1.07	0.00
rs651333	*TULP4*	C/T[b,c,r]	0.34	0.97	1.00	0.00	**0.04[f]**	**1.29**	**0.00**
rs6558002	*EPHX2*	C/T[b,r]	0.24	0.39	1.06	0.00	0.87	1.02	0.00
rs987525	8q24	A/C[b,r]	0.38	**1.22E-03[f]**	**0.81**	**40.55**	0.22	0.86	0.00
rs894673	*FOXE1*	A/T[p]	0.33	0.42	0.95	0.00	0.93	1.01	0.00
rs3758249	*FOXE1*	T/C[p]	0.33	0.56	0.96	0.00	0.90	1.02	0.00
rs7078160	*VAX1*	A/G[r]	0.25	**0.04[f]**	**1.16**	**0.00**	0.88	1.02	0.00
rs4752028	*VAX1*	C/T[b,r]	0.45	0.51	0.96	0.00	0.80	0.97	0.00
rs10785430	*ADAMTS20*	G/A[r]	0.32	0.90	0.99	0.00	0.49	1.09	0.00
rs9574565	*SPRY2*	T/C[b,p]	0.35	0.75	1.02	0.00	0.45	1.10	0.00
rs8001641	*SPRY2*	G/A[b,c,d,p]	0.10	0.35	1.08	0.00	0.37	0.85	0.00
rs17563	*BMP4*	T/C[b,c,d,r]	0.18	0.95	0.99	0.00	0.77	1.04	0.00
rs1258763	*GREM1*	C/T[b,c,d,p]	0.49	0.11	1.11	0.00	0.50	0.92	0.00
rs8049367	*ADCY9*	C/T[c,d,p]	0.30	0.20	1.09	0.00	0.10	0.81	0.00

rs16260	*CDH1*	A/C[r]	0.13	0.59	1.05	0.00	0.39	0.85	0.00
rs11642413	*CDH1*	G/A[b,d,r]	0.28	0.83	1.02	0.00	0.21	0.83	0.00
rs1546124	*CRISPLD2*	G/C[d,r]	0.25	0.60	0.96	0.00	0.89	0.98	0.00
rs4783099	*CRISPLD2*	T/C[r]	0.33	0.59	1.04	0.00	**0.02**[f]	**0.74**	**0.00**
rs8069536	*NTN1*	T/G[r]	0.32	0.13	1.11	0.97	0.88	0.98	0.00
rs8081823	*NTN1*	A/G[p]	0.24	0.08	0.88	0.00	0.63	0.94	32.54
rs17760296	*NOG1*	G/T[r]	0.02	0.92	0.99	0.00	**0.04**[f]	**1.74**	**0.00**
rs227731	*NOG1*	G/T[b,r]	0.22	0.86	0.99	0.00	0.26	1.17	0.00
rs7224837	*AXIN2*	G/A[r]	0.11	0.75	1.04	0.00	0.81	0.95	0.00
rs3923086	*AXIN2*	A/C[b,c,d,r]	0.02	0.25	1.15	0.00	NA	NA	NA
rs17820943	*MAFB*	T/C[p]	0.25	0.33	0.93	15.15	0.68	1.06	22.99
rs13041247	*MAFB*	C/T[p]	0.25	0.37	0.94	34.01	0.42	1.12	0.00
rs11696257	*MAFB*	T/C[p]	0.25	0.30	0.93	32.24	0.61	1.07	0.00
Part B: Meta-analyses of subphenotypes of NSCL/P cohorts from the three countries									
SNP	Probable gene/loci	Minor alleles[a]	African MAF	NSCL			NSCLP		
				p	OR	I	*p*	OR	I
rs1801131	*MTHFR*	C/A[r]	0.15	0.78	1.03	0.00	0.22	1.13	0.00
rs1801133	*MTHFR*	A/G[p]	0.09	0.71	1.06	8.24	0.30	0.30	0.00
rs766325	*PAX7*	G/A[c,d,r]	0.18	0.91	0.99	0.00	0.17	0.86	0.00
rs742071	*PAX7*	T/G[r]	0.39	**0.02**[f]	**1.23**	**68.74**	**0.03**[f]	**1.19**	**0.00**
rs560426	*ABCA4*	C/T[r]	0.49	0.73	1.03	0.00	**0.03**[f]	**1.20**	**10.33**
rs481931	*ABCA4*	T/G[p]	0.10	0.81	0.97	0.00	0.08	1.27	63.75
rs4147811	*ABCA4*	T/C[p]	0.11	0.50	1.10	65.82	0.15	1.21	15.35
rs138751793	*ARHGAP29*	C/T[e]	0.02	0.19	1.53	66.38	0.41	1.29	0.00
rs6677101	*SLC25A24*	G/T[b,d,r]	0.33	0.92	0.99	0.00	0.98	1.00	58.97
rs861020	*IRF6*	A/G[r]	0.11	0.18	1.17	17.72	0.57	1.07	0.00
rs34743335	*IRF6*	T/A	0.02	0.87	0.96	0.00	0.50	0.85	23.72
rs642961	*IRF6*	A/G[r]	0.09	0.96	0.99	15.60	0.15	1.21	62.97
rs7590268	*THADA*	G/T[r]	0.20	0.45	0.92	0.00	0.50	1.07	0.00
rs4332945	*DYSF*	T/G[b,d,r]	0.16	0.54	0.94	10.40	0.71	1.04	0.00
rs2303596	*DYSF*	T/C[c,d,p]	0.22	0.29	0.89	63.58	0.44	0.93	75.54
rs227782	*DYSF*	A/G[b, r]	0.42	0.85	0.98	0.00	0.13	1.14	0.00
rs115200552	*MSX1*	C/G[e]	0.02	0.18	1.37	61.30	0.68	1.10	0.00
rs12532	*MSX1*	G/A[d,p]	0.44	0.55	0.95	0.00	0.51	0.95	0.00
rs2674394	Gene Desert	A/C[r]	0.17	0.06	1.22	0.00	0.42	0.91	0.00
rs651333	*TULP4*	C/T[b,c,r]	0.34	0.63	0.96	0.00	0.74	0.97	0.00
rs6558002	*EPHX2*	C/T[b,r]	0.24	0.82	1.02	0.00	0.11	0.11	0.00
rs987525	8q24	A/C[b,r]	0.38	**5.38E-03**[f]	**1.28**	**0.00**	**0.01**[f]	**0.80**	**54.21**

rs894673	*FOXE1*	A/T[p]	0.33	0.54	0.95	42.39	0.45	0.94	0.00
rs3758249	*FOXE1*	T/C[p]	0.33	0.53	0.94	46.73	0.68	0.96	0.00
rs7078160	*VAX1*	A/G[r]	0.25	**0.03**[f]	**1.23**	**0.00**	0.20	1.13	24.04
rs4752028	*VAX1*	C/T[b,r]	0.45	0.55	1.05	16.64	0.50	0.95	0.00
rs10785430	*ADAMTS20*	G/A[r]	0.32	0.88	1.01	41.30	0.86	0.98	3.00
rs9574565	*SPRY2*	T/C[b,p]	0.35	0.53	1.06	72.62	0.43	1.07	65.44
rs8001641	*SPRY2*	G/A[b,c,d,p]	0.10	0.99	1.00	0.00	0.26	1.13	0.00
rs17563	*BMP4*	A/G[b,c,d,r]	0.18	0.89	0.99	25.84	0.98	1.00	0.00
rs1258763	*GREM1*	C/T[b,c,d,p]	0.49	0.22	0.90	0.00	0.10	1.15	0.00
rs8049367	*ADCY9*	C/T[c,d,p]	0.30	0.36	1.09	10.19	0.35	1.08	0.00
rs16260	*CDH1*	A/C[r]	0.13	0.46	0.91	10.51	0.20	1.16	0.00
rs11642413	*CDH1*	G/A[b,d,r]	0.28	0.98	1.00	0.00	0.55	1.05	0.00
rs1546124	*CRISPLD2*	G/C[d,r]	0.25	0.26	0.90	0.00	0.88	1.01	0.00
rs4783099	*CRISPLD2*	T/C[r]	0.33	0.85	1.02	0.00	0.32	1.09	0.00
rs8069536	*NTN1*	T/G[r]	0.32	0.72	1.03	3.47	**0.04**[f]	**1.20**	**0.00**
rs8081823	*NTN1*	A/G[p]	0.24	0.55	0.95	0.00	0.05	0.83	0.00
rs17760296	*NOG1*	G/T[r]	0.02	0.83	1.04	5.85	0.85	0.97	0.00
rs227731	*NOG1*	G/T[b,r]	0.22	0.38	0.92	0.00	0.59	1.05	0.00
rs7224837	*AXIN2*	G/A[r]	0.11	0.61	1.08	0.00	0.81	1.04	0.00
rs3923086	*AXIN2*	A/C[b,c,d,r]	0.02	0.62	1.10	40.28	NA	NA	0.00
rs17820943	*MAFB*	T/C[p]	0.25	0.25	0.89	15.55	0.43	0.93	0.00
rs13041247	*MAFB*	C/T[p]	0.25	0.25	0.89	31.03	0.54	0.94	0.00
rs11696257	*MAFB*	T/C[p]	0.25	0.24	0.89	27.17	0.40	0.92	0.00

[a]The first allele is the minor allele in Europeans and unless otherwise indicated, the first allele is also the minor allele in Europeans, East Asians, South Asians and Africans, [b]the first allele is the major allele while the second allele is the minor allele in Africans, [c]the first allele is the major allele while the second allele is the minor allele in South Asians, [d]the first allele is the major allele while the second allele is the minor allele in East Asians, [e]first allele is the minor allele and the variation exists only in Africans, [f]loci that reached nominal significance in meta-analyses, [r]minor allele was the risk allele in initial study, [p]minor allele was protective in initial study, **MAF**: minor allele frequency, *p*: *p*-values, OR: odds ratio, I: test of heterogeneity of which 0 to 40 represents no heterogeneity; **NA**: not applicable. All *p*-values reported are for the minor alleles. All initial studies were either carried out in Asians and/or Caucasians, but not Africans. Source of minor alleles and MAF is http://browser.1000genomes.org.

Table 3: Transmission disequilibrium test (TDT) for case-parent trios only

Part A: TDT analyses for NSCL/P and NSCPO							
	Probable	**NSCL/P**			**NSCPO**		
SNP	Gene/Loci	T/NT	*p*	OR (95% CI)	T/NT	*p*	OR (95% CI)
rs1801131	*MTHFR*	27/34	0.37	0.79 (0.48 - 1.32)	10/9	0.82	1.11 (0.45 - 2.73)
rs1801133	*MTHFR*	22/23	0.88	0.96 (0.53 - 1.72)	6/8	0.59	0.75 (0.26 - 2.16)
rs766325	*PAX7*	43/52	0.36	0.83 (0.55 - 1.24)	11/11	1.00	1.00 (0.43 - 2.31)
rs742071	*PAX7*	82/75	0.58	1.09 (0.80 - 1.50)	16/11	0.34	1.46 (0.68 - 3.13)
rs560426	*ABCA4*	78/59	0.10	1.32 (0.94 - 1.85)	18/18	1.00	1.00 (0.52 - 1.92)
rs481931	*ABCA4*	28/25	0.68	1.12 (0.65 - 1.92)	3/8	0.13	0.38 (0.10 - 1.41)
rs4147811	*ABCA4*	26/25	0.89	1.04 (0.60 - 1.80)	5/10	0.20	0.50 (0.17 - 1.46)
rs138751793	*ARHGAP29*	5/7	0.56	0.71 (0.23 - 2.25)	1/2	0.56	0.50 (0.05 - 5.51)
rs6677101	*SLC25A24*	65/75	0.40	0.87 (0.62 - 1.21)	21/14	0.24	1.50 (0.76 - 2.95)
rs861020	*IRF6*	35/29	0.45	1.21 (0.74 - 1.97)	3/7	0.21	0.43 (0.11 - 1.66)
rs34743335	*IRF6*	4/2	0.41	2.00 (0.37 - 10.92)	0/0	NA	NA (NA)
rs642961	*IRF6*	29/29	1.00	1.00 (0.60 - 1.67)	2/7	0.10	0.29 (0.06 - 1.38)
rs7590268	*THADA*	49/48	0.92	1.02 (0.69 - 1.52)	8/8	1.00	1.00 (0.38 - 2.66)
rs4332945	*DYSF*	43/40	0.74	1.08 (0.70 - 1.65)	11/8	0.49	1.38 (0.55 - 3.42)
rs2303596	*DYSF*	45/57	0.23	0.79 (0.53 - 1.18)	12/8	0.37	1.50 (0.61 - 3.67)
rs227782	*DYSF*	73/65	0.50	1.12 (0.80 - 1.57)	20/13	0.22	1.54 (0.77 - 3.09)
rs115200552	*MSX1*	10/13	0.53	0.77 (0.34 - 1.75)	7/2	0.10	3.50 (0.72 - 16.85)
rs12532	*MSX1*	77/71	0.62	1.09 (0.79 - 1.50)	20/22	0.76	0.91 (0.50 - 1.67)
rs2674394	Gene Desert	40/44	0.66	0.91 (0.59 - 1.40)	9/9	1.00	1.00 (0.40 - 2.52)
rs651333	*TULP4*	56/59	0.78	0.95 (0.66 - 1.37)	21/16	0.41	1.31 (0.68 - 2.52)
rs6558002	*EPHX2*	47/40	0.45	1.18 (0.77 - 1.79)	13/12	0.84	1.08 (0.49 - 2.37)
rs987525	8q24	71/59	0.29	1.20 (0.85 - 1.70)	19/20	0.87	0.95 (0.51 - 1.78)
rs894673	*FOXE1*	60/67	0.53	0.90 (0.63 - 1.29)	16/15	0.86	1.07 (0.53 - 2.16)
rs3758249	*FOXE1*	59/66	0.53	0.89 (0.63 - 1.27)	16/15	0.86	1.07 (0.53 - 2.16)
rs7078160	*VAX1*	60/44	0.12	1.36 (0.92 - 2.01)	18/10	0.13	1.80 (0.83 - 3.90)
rs4752028	*VAX1*	73/76	0.81	0.96 (0.70 - 1.32)	**27/13**	**0.03**[b]	**2.08 (1.07 - 4.03)**
rs10785430	*ADAMTS20*	61/59	0.86	1.03 (0.72 - 1.48)	15/11	0.43	1.36 (0.63 - 2.97)
rs9574565	*SPRY2*	69/55	0.21	1.26 (0.88 - 1.79)	18/17	0.87	1.06 (0.55 - 2.05)
rs8001641	*SPRY2*	22/22	1.00	1.00 (0.55 - 1.81)	9/6	0.44	1.50 (0.53 - 4.21)
rs17563	*BMP4*	44/44	1.00	1.00 (0.66 - 1.52)	10/15	0.32	0.67 (0.30 - 1.48)
rs1258763	*GREM1*	73/58	0.19	1.26 (0.89 - 1.78)	19/21	0.75	0.90 (0.49 - 1.68)
rs8049367	*ADCY9*	67/67	1.00	1.00 (0.71 - 1.40)	12/13	0.84	0.92 (0.42 - 2.02)
rs16260	*CDH1*	31/28	0.70	1.11 (0.66 - 1.85)	6/13	0.11	0.46 (0.18 - 1.21)
rs11642413	*CDH1*	62/49	0.22	1.27 (0.87 - 1.84)	14/11	0.55	1.27 (0.58 - 2.80)
rs1546124	*CRISPLD2*	53/44	0.36	1.21 (0.81 - 1.80)	9/14	0.30	0.64 (0.28 - 1.49)
rs4783099	*CRISPLD2*	75/64	0.35	1.17 (0.84 - 1.64)	15/21	0.32	0.71 (0.37 - 1.39)
rs8069536	*NTN1*	67/70	0.80	0.96 (0.68 - 1.34)	14/13	0.85	1.08 (0.51 - 2.29)
rs8081823	*NTN1*	58/56	0.85	1.04 (0.72 - 1.50)	14/15	0.85	0.93 (0.45 - 1.93)
rs17760296	*NOG1*	7/8	0.80	0.88 (0.32 - 2.41)	2/0	0.16	NA (NA)
rs227731	*NOG1*	47/49	0.84	0.96 (0.64 - 1.43)	20/11	0.11	1.82 (0.87 - 3.80)
rs7224837	*AXIN2*	19/27	0.24	0.70 (0.39 - 1.27)	1/6	0.06	0.17 (0.02 - 1.38)

rs3923086	*AXIN2*	2/3	0.65	0.67 (0.11 - 3.99)	1/0	0.32	NA (NA)
rs17820943	*MAFB*	49/42	0.46	1.17 (0.77 - 1.76)	15/12	0.56	1.25 (0.59 - 2.67)
rs13041247	*MAFB*	49/43	0.53	1.14 (0.76 - 1.72)	15/12	0.56	1.25 (0.59 - 2.67)
rs11696257	*MAFB*	48/43	0.60	1.12 (0.74 - 1.69)	14/12	0.69	1.17 (0.54 - 2.52)
Part B: TDT Subphenotype analyses for NSCL/P							
SNP	Probable Gene/Loci	**NSCL** T/NT	*p*	OR (95% CI)	**NSCLP** T/NT	*p*	OR (95% CI)
rs1801131	*MTHFR*	**9/20**	**0.04**[b]	**0.45 (0.20 - 0.99)**	18/14	0.48	1.29 (0.64 - 2.59)
rs1801133	*MTHFR*	7/8	0.80	0.88 (0.31 - 2.41)	15/15	1.00	1.00 (0.49 - 2.05)
rs766325	*PAX7*	18/24	0.35	0.75 (0.41 - 1.38)	25/28	0.68	0.89 (0.52 - 1.53)
rs742071	*PAX7*	**50/30**	**0.03**[b]	**1.67 (1.06 - 2.62)**	32/45	0.14	0.71 (0.45 - 1.12)
rs560426	*ABCA4*	32/35	0.71	0.91 (0.57 - 1.48)	**46/24**	**8.55E-03**[b]	**1.92 (1.17 - 3.14)**
rs481931	*ABCA4*	10/13	0.53	0.77 (0.34 - 1.75)	18/12	0.27	1.50 (0.72 - 3.14)
rs4147811	*ABCA4*	8/10	0.64	0.80 (0.32 - 2.03)	18/15	0.60	1.20 (0.60 - 2.38)
rs138751793	*ARHGAP29*	1/2	0.56	0.50 (0.05 - 5.51)	4/5	0.74	0.80 (0.21 - 2.98)
rs6677101	*SLC25A24*	26/41	0.07	0.63 (0.39 - 1.04)	39/34	0.56	1.15 (0.72 - 1.82)
rs861020	*IRF6*	20/14	0.30	1.43 (0.72 - 2.83)	15/15	1.00	1.00 (0.49 - 2.05)
rs34743335	*IRF6*	2/1	0.56	2.00 (0.18 - 22.06)	2/1	0.56	2.00 (0.18 - 22.06)
rs642961	*IRF6*	16/15	0.86	1.07 (0.53 - 2.16)	13/14	0.85	0.93 (0.44 - 1.98)
rs7590268	*THADA*	21/32	0.13	0.66 (0.38 - 1.14)	28/16	0.07	1.75 (0.95 - 3.23)
rs4332945	*DYSF*	21/17	0.52	1.24 (0.65 - 2.34)	22/23	0.88	0.96 (0.53 - 1.72)
rs2303596	*DYSF*	18/22	0.53	0.82 (0.44 - 1.53)	27/35	0.31	0.77 (0.47 - 1.27)
rs227782	*DYSF*	33/28	0.52	1.18 (0.71 - 1.95)	40/37	0.73	1.08 (0.69 - 1.69)
rs115200552	*MSX1*	6/3	0.32	2.00 (0.50 - 8.00)	4/10	0.11	0.40 (0.13 - 1.28)
rs12532	*MSX1*	39/32	0.41	1.22 (0.76 - 1.95)	38/39	0.91	0.97 (0.62 - 1.52)
rs2674394	Gene Desert	21/17	0.52	1.24 (0.65 - 2.34)	19/27	0.24	0.70 (0.39 - 1.27)
rs651333	*TULP4*	26/26	1.00	1.00 (0.58 - 1.72)	30/33	0.71	0.91 (0.55 - 1.49)
rs6558002	*EPHX2*	15/18	0.60	0.83 (0.42 - 1.65)	32/22	0.17	1.46 (0.85 - 2.50)
rs987525	8q24	35/28	0.38	1.25 (0.76 - 2.06)	36/31	0.54	1.16 (0.72 - 1.88)
rs894673	*FOXE1*	27/31	0.60	0.87 (0.52 - 1.46)	33/36	0.72	0.92 (0.57 - 1.47)
rs3758249	*FOXE1*	27/31	0.60	0.87 (0.52 - 1.46)	32/35	0.71	0.91 (0.57 - 1.48)
rs7078160	*VAX1*	37/23	0.07	1.61 (0.96 - 2.71)	23/21	0.76	1.10 (0.61 - 1.98)
rs4752028	*VAX1*	32/38	0.47	0.84 (0.53 - 1.35)	41/38	0.74	1.08 (0.69 - 1.68)
rs10785430	*ADAMTS20*	25/28	0.68	0.89 (0.52 - 1.53)	36/31	0.54	1.16 (0.72 - 1.88)
rs9574565	*SPRY2*	35/29	0.45	1.21 (0.74 - 1.97)	34/26	0.30	1.31 (0.78 - 2.18)
rs8001641	*SPRY2*	12/12	1.00	1.00 (0.45 - 2.27)	10/10	1.00	1.00 (0.42 - 2.40)
rs17563	*BMP4*	22/16	0.33	1.38 (0.72 - 2.62)	22/28	0.40	0.79 (0.45 - 1.37)
rs1258763	*GREM1*	31/27	0.60	1.15 (0.69 - 1.92)	42/31	0.20	1.36 (0.85 - 2.16)
rs8049367	*ADCY9*	25/28	0.68	0.89 (0.52 - 1.53)	42/39	0.74	1.08 (0.70 - 1.67)
rs16260	*CDH1*	12/14	0.69	0.86 (0.40 - 1.85)	19/14	0.38	1.36 (0.68 - 2.71)
rs11642413	*CDH1*	25/22	0.66	1.14 (0.64 - 2.02)	37/27	0.21	1.37 (0.83 - 2.25)
rs1546124	*CRISPLD2*	25/22	0.66	1.14 (0.61 - 2.02)	28/22	0.40	1.27 (0.73 - 2.23)
rs4783099	*CRISPLD2*	39/35	0.64	1.11 (0.71 - 1.76)	36/29	0.39	1.24 (0.76 - 2.02)
rs8069536	*NTN1*	32/35	0.71	0.91 (0.57 - 1.48)	35/35	1.00	1.00 (0.63 - 1.60)
rs8081823	*NTN1*	30/20	0.16	1.50 (0.85 - 2.64)	28/36	0.32	0.78 (0.47 - 1.27)
rs17760296	*NOG1*	5/2	0.26	2.50 (0.49 - 12.89)	2/6	0.16	0.33 (0.07 - 1.65)

rs227731	*NOG1*	22/26	0.56	0.85 (0.48 - 1.49)	25/23	0.77	1.09 (0.62 - 1.92)
rs7224837	*AXIN2*	10/9	0.82	1.11 (0.45 - 2.73)	9/18	0.08	0.50 (0.22 - 1.11)
rs3923086	*AXIN2*	1/2	0.56	0.50 (0.05 - 5.51)	1/1	1.00	1.00 (0.06 - 15.99)
rs17820943	*MAFB*	18/22	0.53	0.82 (0.44 - 1.53)	31/20	0.12	1.55 (0.88 - 2.72)
rs13041247	*MAFB*	18/22	0.53	0.82 (0.44 - 1.53)	31/21	0.17	1.48 (0.85 - 2.57)
rs11696257	*MAFB*	18/22	0.53	0.82 (0.44 - 1.53)	30/21	0.21	1.43 (0.82 - 2.50)

[b]Loci that demonstrated over-transmission at threshold significance of $p \leq 0.05$, OR: Odds ratio, CI: 95% confidence interval, NA: not applicable.

Table 4: Family-Based Association for Disease Traits (DFAM) for cases and relatives

SNP	Gene/Loci	*p*-values			
		NSCL/P	NSCL	NSCLP	NSCPO
rs1801131	*MTHFR*	0.70	0.68	0.24	0.67
rs1801133	*MTHFR*	0.82	0.51	0.59	0.29
rs766325	*PAX7*	0.61	0.71	0.74	0.24
rs742071	*PAX7*	0.32	**0.02**[b]	0.29	0.96
rs560426	*ABCA4*	**2.59E-02**[b]	0.72	**4.75E-03**[b]	0.80
rs481931	*ABCA4*	0.15	0.55	0.16	0.61
rs4147811	*ABCA4*	0.29	0.44	0.48	0.51
rs138751793	*ARHGAP29*	0.38	0.66	0.43	0.40
rs6677101	*SLC25A24*	1.00	0.80	0.64	0.24
rs861020	*IRF6*	0.43	0.23	0.98	0.35
rs34743335	*IRF6*	0.32	0.52	0.47	0.61
rs642961	*IRF6*	0.83	0.99	0.98	0.15
rs11119388	*SYT14*	0.83	0.85	0.92	0.91
rs7590268	*THADA*	0.85	0.30	0.18	0.77
rs4332945	*DYSF*	**0.04**[b]	**0.02**[b]	0.60	0.62
rs2303596	*DYSF*	0.81	0.84	0.53	0.60
rs227782	*DYSF*	0.36	0.48	0.55	0.47
rs115200552	*MSX1*	0.89	0.13	**3.50E-02**[b]	0.08
rs12532	*MSX1*	0.67	0.96	0.30	0.43
rs2674394	Gene Desert	0.59	0.11	0.58	0.51
rs651333	*TULP4*	0.92	0.90	0.63	0.20
rs6558002	*EPHX2*	0.38	0.77	0.27	0.52
rs987525	8q24	0.80	0.50	0.52	0.99
rs894673	*FOXE1*	0.69	0.88	0.46	0.55
rs3758249	*FOXE1*	0.69	0.86	0.46	0.55
rs7078160	*VAX1*	0.21	0.18	0.77	0.28
rs4752028	*VAX1*	0.88	0.44	0.30	0.06
rs10785430	*ADAMTS20*	0.84	0.86	0.62	0.66
rs9574565	*SPRY2*	0.07	0.16	0.28	0.22
rs8001641	*SPRY2*	0.32	0.19	0.88	0.64
rs375489721	*MIR17HG*	NA	NA	NA	NA
rs185831554	*MIR17HG*	0.32	0.32	NA	NA
rs17563	*BMP4*	0.66	0.15	0.80	0.70
rs1258763	*GREM1*	0.14	1.00	0.06	0.98
rs8049367	*ADCY9*	0.23	0.24	0.56	0.18
rs16260	*CDH1*	0.59	0.59	0.36	0.46
rs11642413	*CDH1*	0.33	0.81	0.08	0.88
rs1546124	*CRISPLD2*	0.30	0.53	0.45	0.15

rs4783099	*CRISPLD2*	0.17	0.14	0.89	0.37
rs8069536	*NTN1*	0.58	0.47	0.87	0.23
rs8081823	*NTN1*	0.97	0.30	0.19	0.89
rs17760296	*NOG1*	0.63	0.25	0.97	0.63
rs227731	*NOG1*	0.24	0.41	0.43	0.09
rs7224837	*AXIN2*	0.20	0.75	0.12	0.35
rs3923086	*AXIN2*	0.89	0.70	**2.88E-03**[b]	0.85
rs17820943	*MAFB*	0.31	0.88	0.14	0.65
rs13041247	*MAFB*	0.37	0.83	0.21	0.63
rs11696257	*MAFB*	0.46	0.89	0.26	0.77

[b]Loci that demonstrated over-transmission at threshold significance, **NA**: not applicable.

Table 5: Novel, rare and potentially aetiologic variants observed in sequenced genes

Part A: Variants observed in cases and some parents but not in controls				
HGVS	HGVp	Total number of cases with variant	Subphenotype of cases with variant	Segregation analyses
ARHGAP29				
c.341-30T>A	N/A	1	NSCL	N/A
c.511-107T>C	N/A	2	NSCLP and NSCPO	N/A
c.967A>G	p.Asn323Asp	1	NSCL	Absent in father
c.1277delAinsTA	p.Lys426IlefsTer6	1	NSCLP	Absent in mother
c.1281+4A>G	N/A	1	NSCLP	Observed in clinically unaffected mother
PAX7				
c.1227G>A	p.Leu409Leu	1	NSCL	N/A
Part B: Bioinformatics-predicted effects of potentially pathogenic variants				
HGVS	Polyphen-2	SIFT	Human Splice Finder	RegulomeDB
ARHGAP29				
c.341-30T>A	N/A	N/A	Alteration of ESS site	N/A
c.511-107T>C	N/A	N/A	Alteration of ESS site and creation of new ESE site	N/A
c.967A>G	Benign	Deleterious	N/A	N/A
c.1277delAinsTA	N/A	N/A	N/A	N/A
c.1281+4A>G	N/A	N/A	Alteration of wildtype donor site	N/A
PAX7				
c.1227G>A	Benign	Tolerated	Alteration of an ESE site	N/A

ESS: Exonic Splicing Silencer, **ESE:** Exonic Splicing Enhancer, **N/A:** Not Applicable, **NSCLP**: nonsyndromic cleft lip and palate, **NSCL**: nonsyndromic cleft lip only, **NSCPO**: nonsyndromic cleft palate only. All analyses were based on genome assembly number GRCh37/hg19, 2009 (http://genome.ucsc.edu).

Supplemental Methods

Eligible subjects or participants

Eligible subjects were individuals with NSOFCs and their families, born to indigenous Ghanaian, Ethiopian and Nigerian parents. These families were recruited at the cleft clinics and during surgical missions. Births from Caucasians and Asians were excluded. Controls were recruited in Ghana, Nigeria and Ethiopia at the immunization clinics and dental clinics to match cases recruited from each of these countries. Controls were Africans born alive without any congenital birth defects in Ghana, Ethiopia and Nigeria. In Nigeria, two different centers (Lagos and Ife) coordinated patient recruitment. Only one center each coordinated patient recruitment in Ghana and Ethiopia. We have previously described individuals that are involved in recruitments for our cleft studies in Africa (Butali et al. 2011; Butali et al. 2015). In summary, recruitment is done by surgeons (i.e. plastic surgery, ear nose and throat surgeons, pediatric surgeons, maxillofacial surgeons and dental surgeons).

DNA Collection and processing

We collected saliva and cheek swab samples from participants using Oragene DNA Collection Kits (http://www.dnagenotek.com). We extracted DNA from both saliva and cheek swab samples using the Oragene Saliva processing protocol (http://genetics.uiowa.edu/protocols.php). We then determined the concentration of DNA using Quibit Assay that employed Quibit 2.0 Fluorometer (http://www.invitrogen.com/site/us/en/home/brands/Product-Brand/Qubit.html). We finally performed XY-Genotyping on all samples to validate the sexes and sanctity of the samples (http://genetics.uiowa.edu/protocols.php).

SNP Genotyping

The detailed protocol is available at Murray Laboratory (http://genetics.uiowa.edu/protocols.php) but a summary is presented here. We selected these SNPs based on GWAS and candidate gene studies. We randomly assigned each sample to a well in a labeled 96-well microplate to form a Plate Map, using sample concentration of 2ng/ul. Each of these microplates also contained two template controls, NA18856 (male) and NA18855 (female). These two template controls are Yoruba HapMap samples. They therefore served as a guide in calling the genotype of individuals genotyped in this study. Each microplate also had provision for at least two No Template Controls (NTCs), which was dH_2O; however, NTCs were not added unto Microplates until the running of the chips. SNPs were designed based on human genome assembly GRCh37/hg19, 2009 (http://genome.ucsc.edu) and were obtained from ABi/Life Technologies (www.lifetechnologies.com).

DNA sequencing and DNA sequence analyses

The protocols for primer design and optimization as well as DNA amplification by PCR and electrophoresis have been described earlier (Butali et al. 2014). We shipped PCR products to Functional Biosciences, Madison, Wisconsin (http://order.functionalbio.com/seq/index) where they were sequenced using an ABI 3730XL (http://www.appliedbiosystems.com/absite/us/en/home.html). Chromatograms were then transferred to a Unix workstation, base-called with PHRED (http://www.phrap.org/phredphrapconsed.html, v.0.961028), assembled with PHRAP (http://www.phrap.org/, v.0.960731), scanned by POLYPHRED (http://droog.gs.washington.edu/polyphred/, v. 0.970312) and viewed with CONSED programme (http://www.phrap.org/consed/consed.html, v. 4.0).

We ascertained the genomic location of each variant revealed by CONSED by employing the “Blat” function of UCSC Genome Browser (https://genome.ucsc.edu/). We predicted the functional effect of a coding variant on protein using Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/), SIFT (http://sift.jcvi.org/) and Ensemble (http://www.ensembl.org/Homo_sapiens/Tools/VEP). Effect of a variant on mRNA splicing was ascertained using Human Splicing Finder 3.0 (http://www.umd.be/HSF3/). Finally, we predicted the effect of a mutation on a regulatory region using RegulomeDB (http://regulomedb.org/).

We ascertained the Minor Allele Frequencies (MAF) or novelty of a mutation by comparing it to variants in 1000 Genomes (http://browser.1000genomes.org/index.html), Exome Variant Server (http://evs.gs.washington.edu/EVS/), dbSNP (www.ncbi.nlm.nih.gov/SNP/), ExAC Browser (http://exac.broadinstitute.org/) and other literature on OFCs. We classified mutations as “novel” if they have never been reported in any of these databases or literature.

Table S1: List of 48 SNPs that were genotyped

Chromosome	coordinate	SNP	Probable gene/loci	Alleles/ Variation	NA18856 Genotype in 1000 Genomes	NA18855 Genotype in 1000 Genomes	Averge Call Rate (%)	Reference study	Study population
1	11854476	rs1801131	*MTHFR*	T>G	T/T	T/T	99.4	Boyles et al.2008	Europeans
1	11856378	rs1801133	*MTHFR*	G>A	G/A	G/G	99.7	Boyles et al.2008	Europeans
1	18956458	rs766325	*PAX7*	A>G	A/A	A/A	99.6	Beaty et al. 2010	Europeans and Asians
1	18979874	rs742071	*PAX7*	G>T	T/G	G/G	99.4	Beaty et al. 2010	Europeans and Asians
1	94553438	rs560426	*ABCA4*	T>C	C/T	T/C	99.5	Beaty et al. 2010	Europeans and Asians
1	94570016	rs481931	*ABCA4*	G>T	G/G	No data	99.5	Beaty et al. 2010	Europeans and Asians
1	94575056	rs4147811	*ABCA4*	C>T	C/C	No data	99.3	Beaty et al. 2010	Europeans and Asians
1	94650805	rs138751793	*ARHGAP29*	T>C	T/T	No data	99.3	Present Study	Africans
1	108699730	rs6677101	*SLC25A24*	T>G	G/T	G/G	99.1	Butler et al. 2015	Europeans
1	209977111	rs861020	*IRF6*	G>A	G/G	G/G	99.5	Rojas-Martinez et al.2010	Europeans
1	209979529	rs34743335	*IRF6*	A>T	A/A	No data	97.9	Pegelow et al. 2008	Europeans
1	209989270	rs642961	*IRF6*	G>A	G/G	G/G	99.5	Rahimov et al. 2008	Europeans and Asians
1	210174417	rs11119388	*SYT14*	A>G	A/A	A/A	99.6	Leslie et al. 2014	Europeans and Asians
2	43540125	rs7590268	*THADA*	T>G	T/T	T/T	99.5	Mangold et al. 2010	Europeans
2	71674476	rs4332945	*DYSF*	T>G	T/T	T/T	99.3	Brayton et al. 2009	Mouse screen
2	71780215	rs2303596	*DYSF*	C>T	C/T	C/C	99.2	Brayton et al. 2009	Mouse screen
2	71866842	rs227782	*DYSF*	A>G	A/G	A/G	99.3	Brayton et al. 2009	Mouse screen

4	4865146	rs12532	*MSX1*	A>G	G/A	A/G	99.4	Suzuki et al. 2004	Asians
4	4864991	rs115200552	*MSX1*	G>C	G/G	No data	99.5	Present study	Africans
6	93506409	rs2674394	Gene Desert	A>C	C/C	C/C	99.4	Ludwig et al. 2012	Europeans and Asians
6	158885758	rs651333	*TULP4*	C>T	C/T	T/T	99.3	Ludwig et al. 2012	Europeans and Asians
8	27389542	rs6558002	*EPHX2*	C>T	C/C	C/C	99.1	Ludwig et al. 2012	Europeans and Asians
8	129946154	rs987525	8q24	A>C	A/A	C/C	99.4	Birnbaum et al. 2009	Europeans
9	100612270	rs894673	*FOXE1*	A>T	T/T	T/T	99.3	Moreno et al. 2009	Europeans, Asians and Hispanics
9	100614140	rs3758249	*FOXE1*	T>C	C/C	C/C	99.4	Moreno et al. 2009	Europeans, Asians and Hispanics
10	118827560	rs7078160	*VAX1*	G>A	A/G	A/G	99.2	Beaty et al. 2010	Europeans and Asians
10	118834991	rs4752028	*VAX1*	T>C	C/C	C/T	99.6	Beaty et al. 2010	Europeans and Asians
12	43819298	rs10785430	*ADAMTS20*	A>G	A/A	A/A	99.5	Wolf et al. 2015	Hispanics
13	80668874	rs9574565	*SPRY2*	T>C	C/T	T/T	99.4	Ludwig et al. 2012	Europeans and Asians
13	80692811	rs8001641	*SPRY2*	G>A	A/G	G/G	99.5	Ludwig et al. 2012	Europeans and Asians
13	92003297	rs375489721	*MIR17HG*	T>C	T/T	No data	99.1	Amendt et al. unplished	Mouse screen
13	92003356	rs185831554	*MIR17HG*	T>G	T/T	No data	99.3	Amendt et al. unplished	Mouse screen
14	54417522	rs17563	*BMP4*	A>G	A/A	G/A	99.5	Chen et al. 2008	Asians
15	33050423	rs1258763	*GREM1*	C>T	C/C	C/C	99.3	Ludwig et al. 2012	Europeans and Asians
16	3980445	rs8049367	*ADCY9*	C>T	C/T	C/T	99.4	Sun et al. 2015	Asians
16	68771034	rs16260	*CDH1*	C>A	C/A	A/C	99.2	Song et al. 2011	Asians
16	68790394	rs11642413	*CDH1*	G>A	G/G	A/G	99.5	Song et al. 2011	Asians

16	84872051	rs1546124	*CRISPLD2*	C>G	C/C	C/G	99.5	Chiquet et al. 2007	Hispanics
16	84941329	rs4783099	*CRISPLD2*	C>T	C/T	C/C	99.3	Chiquet et al. 2007	Hispanics
17	8956285	rs8069536	*NTN1*	G>T	G/G	G/T	99.4	Beaty et al. 2010	Europeans and Asians
17	8965551	rs8081823	*NTN1*	G>A	A/G	G/G	99.4	Beaty et al. 2010	Europeans and Asians
17	54615617	rs17760296	*NOG1*	T>G	T/T	T/T	99.5	Mangold et al. 2010	Europeans
17	54773238	rs227731	*NOG1*	G>T	T/G	T/G	99.3	Mangold et al. 2010	Europeans
17	63528123	rs7224837	*AXIN2*	A>G	A/G	A/A	99.3	Letra et al. 2012	Europeans and Asians
17	63549488	rs3923086	*AXIN2*	A>C	A/A	A/A	99.7	Letra et al. 2012	Europeans and Asians
20	39268516	rs17820943	*MAFB*	C>T	T/C	C/C	99.6	Beaty et al. 2010	Europeans and Asians
20	39269074	rs13041247	*MAFB*	T>C	C/T	T/T	99.5	Beaty et al. 2010	Europeans and Asians
20	39270816	rs11696257	*MAFB*	C>T	T/C	C/C	99.3	Beaty et al. 2010	Europeans and Asians

Note: Except the studies designated as “present study” and Moreno et al. 2009, these loci were largely associated with NSCL/P in the study populations.

Table S2: Case-control analyses for Ethiopia

Part A: Case-control analyses for NSCL/P and NSCPO for Ethiopia							
SNP	Probable gene/loci	NSCL/P			NSCPO		
		p	OR	95% CI	*p*	OR	95% CI
rs1801131	*MTHFR*	0.37	1.10	0.89 - 1.36	0.88	0.95	0.49 - 1.85
rs1801133	*MTHFR*	0.85	0.97	0.71 - 1.34	0.83	0.89	0.31 - 2.54
rs766325	*PAX7*	0.39	0.90	0.71 - 1.14	0.86	0.94	0.46 - 1.92
rs742071	*PAX7*	**5.57E-03[a]**	**1.33**	**1.09 - 1.63**	0.53	0.82	0.44 - 1.53
rs560426	*ABCA4*	0.95	0.99	0.81 - 1.22	0.23	1.46	0.79 - 2.71
rs481931	*ABCA4*	0.75	0.95	0.68 - 1.32	**0.03[a]**	**0.00**	**0.00 - NA**
rs4147811	*ABCA4*	0.69	0.94	0.67 - 1.30	**0.03[a]**	**0.00**	**0.00 - NA**
rs138751793	*ARHGAP29*	0.62	0.59	0.07 - 4.88	0.05	6.42	0.73 - 56.15
rs6677101	*SLC25A24*	0.30	0.89	0.72 - 1.11	0.52	1.23	0.66 - 2.29
rs861020	*IRF6*	0.11	1.21	0.96 - 1.52	0.13	1.66	0.86 - 3.19
rs34743335	*IRF6*	0.27	0.78	0.51 - 1.21	0.52	0.63	0.15 - 2.63
rs642961	*IRF6*	**0.02[a]**	**1.44**	**1.07 - 1.94**	0.22	1.68	0.73 - 3.84
rs7590268	*THADA*	0.47	0.92	0.73 - 1.16	0.70	1.14	0.58 - 2.26
rs4332945	*DYSF*	0.45	0.92	0.74 - 1.14	0.77	0.91	0.47 - 1.74
rs2303596	*DYSF*	**2.31E-03[a]**	**0.69**	**0.54 - 0.87**	0.10	0.51	0.22 - 1.15
rs227782	*DYSF*	0.10	0.84	0.68 - 1.04	0.09	0.55	0.27 - 1.10
rs115200552	*MSX1*	0.40	1.46	0.60 - 3.55	0.33	2.66	0.34 - 20.96
rs12532	*MSX1*	0.59	1.06	0.86 - 1.31	0.22	0.65	0.32 - 1.31
rs2674394	Gene Desert	0.69	0.95	0.74 - 1.23	0.73	1.14	0.54 - 2.41
rs651333	*TULP4*	0.52	0.94	0.76 - 1.15	0.70	0.88	0.47 - 1.67
rs6558002	*EPHX2*	0.44	0.92	0.76 - 1.13	0.44	0.78	0.42 - 1.46
rs987525	8q24	**7.82E-04[a]**	**1.41**	**1.15 - 1.73**	0.20	1.50	0.81 - 2.78
rs894673	*FOXE1*	0.47	0.93	0.75 - 1.15	0.53	1.23	0.65 - 2.29
rs3758249	*FOXE1*	0.52	0.93	0.75 - 1.16	0.52	1.23	0.66 - 2.30
rs7078160	*VAX1*	0.48	1.10	0.85 - 1.41	0.56	0.77	0.32 - 1.86
rs4752028	*VAX1*	0.90	0.99	0.81 - 1.21	0.50	0.80	0.42 - 1.52
rs10785430	*ADAMTS20*	0.83	1.02	0.82 - 1.27	0.69	1.14	0.59 - 2.19
rs9574565	*SPRY2*	0.87	0.98	0.80 - 1.21	0.57	0.83	0.43 - 1.59
rs8001641	*SPRY2*	0.21	1.15	0.92 - 1.43	0.19	0.59	0.27 - 1.30
rs17563	*BMP4*	0.91	0.99	0.80 - 1.22	0.75	0.90	0.46 - 1.75
rs1258763	*GREM1*	0.15	0.86	0.70 - 1.06	0.79	1.09	0.59 - 2.02
rs8049367	*ADCY9*	0.95	0.99	0.80 - 1.23	0.36	0.72	0.36 - 1.46
rs16260	*CDH1*	0.46	1.11	0.84 - 1.48	0.66	0.81	0.31 - 2.08
rs11642413	*CDH1*	0.39	1.09	0.89 - 1.34	0.84	1.07	0.58 - 1.97
rs1546124	*CRISPLD2*	0.44	0.92	0.75 - 1.13	0.36	0.74	0.39 - 1.41
rs4783099	*CRISPLD2*	0.87	0.98	0.80 - 1.21	0.17	0.62	0.32 - 1.23

rs8069536	*NTN1*	0.77	1.04	0.81 - 1.32	0.15	0.51	0.20 - 1.31
rs8081823	*NTN1*	0.20	0.87	0.69 - 1.08	**0.04**[a]	**0.41**	**0.17 - 0.98**
rs17760296	*NOG1*	0.48	0.91	0.70 - 1.18	0.51	1.28	0.62 - 2.63
rs227731	*NOG1*	0.45	0.93	0.76 - 1.13	0.71	0.89	0.48 - 1.65
rs7224837	*AXIN2*	0.53	1.16	0.73 - 1.85	0.16	0.00	0.00 - NA
rs3923086	*AXIN2*	0.42	1.11	0.87 - 1.41	0.59	1.22	0.59 - 2.52
rs17820943	*MAFB*	0.06	0.81	0.65 - 1.01	0.27	0.68	0.34 - 1.36
rs13041247	*MAFB*	**0.04**[a]	**0.80**	**0.64 - 0.99**	0.51	0.80	0.41 - 1.55
rs11696257	*MAFB*	**0.04**[a]	**0.79**	**0.64 - 0.99**	0.51	0.80	0.41 - 1.55
Part B: Case-control analyses for NSCL/P subphenotypes for Ethiopia							
SNP	Probable gene/loci	NSCL			NSCLP		
		p	OR	95% CI	*p*	OR	95% CI
rs1801131	*MTHFR*	0.37	1.15	0.85 - 1.57	0.57	1.08	0.82 - 1.42
rs1801133	*MTHFR*	0.42	0.81	0.49 - 1.34	0.58	1.12	0.75 - 1.66
rs766325	*PAX7*	0.68	0.93	0.66 - 1.31	0.56	0.91	0.68 - .123
rs742071	*PAX7*	**7.74E-03**[a]	**1.49**	**1.11 - 2.00**	0.06	1.28	0.99 - 1.65
rs560426	*ABCA4*	0.42	0.88	0.66 - 1.19	0.57	1.08	0.83 - 1.40
rs481931	*ABCA4*	0.70	0.91	0.56 - 1.49	0.91	1.03	0.67 - 1.56
rs4147811	*ABCA4*	0.62	0.88	0.54 - 1.44	0.96	1.01	0.66 - 1.53
rs138751793	*ARHGAP29*	0.45	0.00	0.00 - NA	0.94	0.93	0.11 - 7.76
rs6677101	*SLC25A24*	0.93	0.99	0.72 - 1.34	0.30	0.87	0.66 - 1.14
rs861020	*IRF6*	0.07	1.35	0.98 - 1.87	0.31	1.17	0.87 - 1.57
rs34743335	*IRF6*	0.53	0.82	0.44 - 1.53	0.26	0.73	0.42 - 1.27
rs642961	*IRF6*	0.49	1.18	0.74 - 1.86	**9.11E-03**[a]	**1.61**	**1.12 - 2.31**
rs7590268	*THADA*	0.17	0.78	0.54 - 1.11	0.98	1.00	0.75 - 1.34
rs4332945	*DYSF*	0.14	0.79	0.57 - 1.08	0.96	1.01	0.77 - 1.32
rs2303596	*DYSF*	**4.99E-03**[a]	**0.59**	**0.41 - 0.86**	**0.03**[a]	**0.70**	**0.52 - 0.96**
rs227782	*DYSF*	0.41	0.88	0.65 - 1.19	0.17	0.83	0.63 - 1.09
rs115200552	*MSX1*	**0.04**[a]	**2.81**	**0.99 - 7.97**	0.69	0.74	0.17 - 3.26
rs12532	*MSX1*	0.82	1.04	0.76 - 1.41	0.58	1.08	0.83 - 1.41
rs2674394	Gene Desert	0.70	1.07	0.75 - 1.54	0.60	0.91	0.66 - 1.27
rs651333	*TULP4*	0.35	0.86	0.64 - 1.17	0.79	0.96	0.74 - 1.25
rs6558002	*EPHX2*	0.79	0.96	0.72 - 1.29	0.35	0.88	0.68 - 1.15
rs987525	8q24	**0.03**[a]	**1.38**	**1.03 - 1.85**	**1.07E-03**[a]	**1.54**	**1.19 - 1.99**
rs894673	*FOXE1*	0.17	0.80	0.58 - 1.10	0.99	1.00	0.77 - 1.31
rs3758249	*FOXE1*	0.12	0.78	0.56 - 1.07	0.79	1.04	0.79 - 1.36
rs7078160	*VAX1*	0.87	0.97	0.66 - 1.43	0.20	1.23	0.89 - 1.68
rs4752028	*VAX1*	0.85	0.97	0.72 - 1.31	0.99	1.00	0.77 - 1.30
rs10785430	*ADAMTS20*	0.14	1.26	0.93 - 1.72	0.36	0.87	0.66 - 1.16
rs9574565	*SPRY2*	**7.05E-03**[a]	**1.50**	**1.11 - 2.01**	**0.02**[a]	**0.73**	**0.55 - 0.95**
rs8001641	*SPRY2*	0.72	0.94	0.68 - 1.31	0.09	1.27	0.97 - 1.67

rs17563	*BMP4*	0.93	0.99	0.72 - 1.35	0.91	1.02	0.77 - 1.33
rs1258763	*GREM1*	0.44	0.89	0.66 - 1.20	0.15	0.82	0.63 - 1.07
rs8049367	*ADCY9*	0.69	0.94	0.68 - 1.29	0.81	1.03	0.79 - 1.36
rs16260	*CDH1*	0.83	1.05	0.69 - 1.59	0.58	1.11	0.77 - 1.59
rs11642413	*CDH1*	0.27	1.18	0.88 - 1.58	0.66	1.06	0.82 - 1.37
rs1546124	*CRISPLD2*	0.47	0.90	0.66 - 1.21	0.62	0.94	0.72 - 1.21
rs4783099	*CRISPLD2*	0.49	0.90	0.66 - 1.22	0.72	1.05	0.81 - 1.36
rs8069536	*NTN1*	0.65	0.92	0.64 - 1.33	0.28	1.18	0.87 - 1.59
rs8081823	*NTN1*	0.84	0.97	0.70 - 1.33	0.30	0.86	0.64 - 1.15
rs17760296	*NOG1*	0.42	0.85	0.58 - 1.26	0.68	0.93	0.67 - 1.30
rs227731	*NOG1*	0.19	0.82	0.61 - 1.10	0.84	0.97	0.75 - 1.26
rs7224837	*AXIN2*	0.61	1.19	0.62 - 2.29	0.54	1.20	0.67 - 2.18
rs3923086	*AXIN2*	0.79	1.05	0.73 - 1.50	0.19	1.22	0.90 - 1.66
rs17820943	*MAFB*	**0.02**[a]	**0.68**	**0.48 - 0.95**	0.24	0.85	0.64 - 1.12
rs13041247	*MAFB*	**0.02**[a]	**0.67**	**0.48 - 0.93**	0.22	0.84	0.64 - 1.11
rs11696257	*MAFB*	**0.02**[a]	**0.67**	**0.48 - 0.93**	0.19	0.83	0.63 - 1.09

[a]Loci that reached nominal significance

Table S3: Case-control analyses for Ghana

Part A: Case-control analyses for NSCL/P and NSCPO for Ghana							
	Probable	NSCL/P			NSCPO		
SNP	gene/loci	*p*	OR	95% CI	*p*	OR	95% CI
rs1801131	*MTHFR*	0.82	1.03	0.79 - 1.35	0.40	0.81	0.51 - 1.32
rs1801133	*MTHFR*	0.94	1.01	0.72 - 1.43	0.51	0.81	0.44 - 1.49
rs766325	*PAX7*	0.82	0.97	0.75 - 1.26	0.65	0.91	0.60 - 1.38
rs742071	*PAX7*	0.84	1.02	0.85 - 1.23	0.88	0.98	0.72 - 1.33
rs560426	*ABCA4*	**0.03**[a]	**1.22**	**1.01 - 1.47**	0.16	1.24	0.92 - 1.67
rs481931	*ABCA4*	0.67	1.07	0.78 - 1.49	0.47	0.81	0.46 - 1.43
rs4147811	*ABCA4*	0.89	1.02	0.74 - 1.42	0.64	0.88	0.50 - 1.52
rs138751793	*ARHGAP29*	0.77	1.10	0.57 - 2.12	0.68	0.77	0.23 - 2.63
rs6677101	*SLC25A24*	0.33	1.10	0.91 - 1.34	0.27	1.19	0.87 - 1.63
rs861020	*IRF6*	0.88	0.98	0.72 - 1.32	0.27	0.73	0.41 - 1.28
rs34743335	*IRF6*	0.93	1.04	0.43 - 2.50	0.52	0.52	0.07 - 4.07
rs642961	*IRF6*	0.48	0.89	0.63 - 1.24	0.13	0.60	0.30 - 1.17
rs7590268	*THADA*	0.80	1.03	0.82 - 1.30	0.26	0.79	0.52 - 1.20
rs4332945	*DYSF*	0.87	1.02	0.79 - 1.32	0.70	0.92	0.61 - 1.40
rs2303596	*DYSF*	0.99	1.00	0.80 - 1.25	0.05	1.40	1.00 - 1.98
rs227782	*DYSF*	0.75	1.03	0.85 - 1.25	0.30	1.17	0.87 - 1.59
rs115200552	*MSX1*	0.24	1.28	0.85 - 1.91	0.06	1.72	0.97 - 3.03
rs12532	*MSX1*	0.33	0.91	0.76 - 1.10	0.87	1.03	0.76 - 1.38
rs2674394	Gene Desert	0.55	1.08	0.84 - 1.38	0.75	1.07	0.71 - 1.60
rs651333	*TULP4*	0.74	0.97	0.79 - 1.18	0.14	1.27	0.93 - 1.73
rs6558002	*EPHX2*	0.52	1.08	0.86 - 1.35	0.67	0.92	0.62 - 1.36
rs987525	8q24	0.11	0.85	0.70 - 1.04	0.80	0.96	0.70 - 1.31
rs894673	*FOXE1*	0.24	0.89	0.73 - 1.08	0.56	0.91	0.66 - 1.25
rs3758249	*FOXE1*	0.30	0.90	0.74 - 1.10	0.59	0.92	0.67 - 1.26
rs7078160	*VAX1*	**0.03**[a]	**1.25**	**1.02 - 1.54**	0.59	1.10	0.78 - 1.55
rs4752028	*VAX1*	0.12	0.86	0.72 - 1.04	0.54	0.91	0.67 - 1.23
rs10785430	*ADAMTS20*	0.46	0.93	0.75 - 1.14	0.20	1.23	0.89 - 1.70
rs9574565	*SPRY2*	0.45	1.08	0.89 - 1.30	0.47	1.12	0.83 - 1.51
rs8001641	*SPRY2*	0.79	1.04	0.78 - 1.40	0.55	0.86	0.53 - 1.41
rs17563	*BMP4*	0.45	0.91	0.72 - 1.16	0.99	1.00	0.69 - 1.45
rs1258763	*GREM1*	0.63	1.05	0.87 - 1.26	0.65	0.93	0.69 - 1.26
rs8049367	*ADCY9*	0.34	1.10	0.90 - 1.34	0.48	0.89	0.64 - 1.23
rs16260	*CDH1*	0.84	1.03	0.78 - 1.37	0.46	0.84	0.52 - 1.35
rs11642413	*CDH1*	0.81	0.97	0.78 - 1.21	0.24	0.80	0.55 - 1.16
rs1546124	*CRISPLD2*	0.92	1.01	0.80 - 1.28	0.44	1.15	0.80 - 1.65
rs4783099	*CRISPLD2*	0.41	1.08	0.90 - 1.31	0.28	0.84	0.61 - 1.15
rs8069536	*NTN1*	0.70	1.04	0.86 - 1.27	0.88	0.98	0.71 - 1.34

rs8081823	*NTN1*	0.46	0.92	0.75 - 1.14	0.73	1.06	0.76 - 1.47
rs17760296	*NOG1*	0.90	1.05	0.48 - 2.28	0.07	2.66	0.90 - 7.87
rs227731	*NOG1*	0.43	1.10	0.87 - 1.38	0.25	1.23	0.86 - 1.75
rs7224837	*AXIN2*	0.80	1.04	0.75 - 1.44	0.49	0.81	0.44 - 1.48
rs3923086	*AXIN2*	0.36	1.99	0.44 - 8.92	0.33	0.00	0.00 - NA
rs17820943	*MAFB*	0.88	1.02	0.81 - 1.28	0.82	1.04	0.73 - 1.49
rs13041247	*MAFB*	0.70	1.05	0.83 - 1.32	0.58	1.11	0.77 - 1.59
rs11696257	*MAFB*	0.84	1.02	0.81 - 1.29	0.88	1.03	0.72 - 1.48
Part B: Case-control analyses for NSCL/P subphenotypes for Ghana							
SNP	Probable gene/loci	NSCL			NSCLP		
		p	OR	95% CI	*p*	OR	95% CI
rs1801131	*MTHFR*	0.58	0.90	0.63 - 1.30	0.41	1.17	0.81 - 1.68
rs1801133	*MTHFR*	0.91	1.03	0.64 - 1.64	0.92	1.02	0.64 - 1.63
rs766325	*PAX7*	0.57	1.10	0.79 - 1.53	0.30	0.82	0.57 - 1.19
rs742071	*PAX7*	0.83	0.97	0.76 - 1.25	0.48	1.10	0.85 - 1.42
rs560426	*ABCA4*	0.38	1.12	0.87 - 1.42	**0.01**[a]	**1.39**	**1.08 - 1.80**
rs481931	*ABCA4*	0.82	1.05	0.68 - 1.62	0.76	1.07	0.69 - 1.67
rs4147811	*ABCA4*	0.57	0.88	0.55 - 1.39	0.47	1.17	0.76 - 1.80
rs138751793	*ARHGAP29*	0.54	0.71	0.24 - 2.09	0.36	1.43	0.66 - 3.09
rs6677101	*SLC25A24*	0.98	1.00	0.77 - 1.31	0.09	1.26	0.97 - 1.65
rs861020	*IRF6*	0.73	1.07	0.73 - 1.57	0.62	0.90	0.58 -1 .38
rs34743335	*IRF6*	0.89	1.08	0.35 - 3.30	0.87	0.90	0.26 - 3.15
rs642961	*IRF6*	0.89	0.97	0.63 - 1.50	0.39	0.81	0.50 - 1.32
rs7590268	*THADA*	0.86	0.97	0.71 - 1.33	0.25	1.20	0.88 - 1.64
rs4332945	*DYSF*	0.88	1.03	0.73 - 1.43	0.82	1.04	0.73 - 1.48
rs2303596	*DYSF*	0.46	1.12	0.83 - 1.50	0.69	0.94	0.68 - 1.29
rs227782	*DYSF*	0.55	0.93	0.72 - 1.19	0.20	1.18	0.91 - 1.53
rs115200552	*MSX1*	0.33	1.31	0.76 - 2.27	0.39	1.27	0.73 - 2.18
rs12532	*MSX1*	0.68	0.95	0.74 - 1.21	0.16	0.83	0.64 - 1.08
rs2674394	Gene Desert	0.08	1.32	0.97 - 1.80	0.46	0.87	0.60 - 1.26
rs651333	*TULP4*	0.72	1.05	0.81 - 1.36	0.44	0.90	0.68 - 1.18
rs6558002	*EPHX2*	0.78	0.96	0.70 - 1.30	0.14	1.25	0.93 - 1.69
rs987525	8q24	0.06	0.78	0.60 - 1.02	0.56	0.92	0.70 - 1.21
rs894673	*FOXE1*	0.43	0.90	0.70 - 1.17	0.31	0.87	0.66 - 1.14
rs3758249	*FOXE1*	0.46	0.91	0.70 - 1.17	0.41	0.89	0.68 - 1.17
rs7078160	*VAX1*	**0.04**[a]	**1.32**	**1.02- 1.72**	0.17	1.22	0.92 - 1.63
rs4752028	*VAX1*	0.11	0.82	0.64 - 1.05	0.27	0.87	0.67 - 1.12
rs10785430	*ADAMTS20*	0.48	0.91	0.69 - 1.19	0.83	0.97	0.73 - 1.29
rs9574565	*SPRY2*	0.19	1.18	0.92 - 1.50	0.70	0.95	0.73 - 1.24
rs8001641	*SPRY2*	0.65	1.09	0.75 - 1.59	0.75	0.93	0.62 - 1.41
rs17563	*BMP4*	0.49	0.89	0.65 - 1.23	0.61	0.92	0.66 - 1.28

rs1258763	*GREM1*	0.45	1.10	0.86 - 1.40	0.65	1.06	0.82 - 1.37
rs8049367	*ADCY9*	0.56	1.08	0.83 - 1.40	0.49	1.10	0.84 - 1.44
rs16260	*CDH1*	0.10	0.70	0.46 - 1.07	0.12	1.33	0.93 - 1.91
rs11642413	*CDH1*	0.85	0.97	0.73 - 1.30	1.00	1.00	0.74 - 1.35
rs1546124	*CRISPLD2*	0.49	0.89	0.65 - 1.23	0.30	1.18	0.86 - 1.61
rs4783099	*CRISPLD2*	0.78	1.04	0.81 - 1.33	0.17	1.20	0.92 - 1.56
rs8069536	*NTN1*	0.66	0.94	0.72 - 1.23	0.31	1.15	0.88 - 1.50
rs8081823	*NTN1*	0.80	0.96	0.73 - 1.27	0.33	0.86	0.64 - 1.16
rs17760296	*NOG1*	0.71	1.21	0.44 - 3.32	0.66	1.29	0.42 - 3.99
rs227731	*NOG1*	0.61	1.08	0.80 - 1.46	0.53	1.11	0.81 - 1.52
rs7224837	*AXIN2*	0.62	1.11	0.73 - 1.69	0.80	1.06	0.67 - 1.68
rs3923086	*AXIN2*	0.15	2.86	0.64 - 12.83	0.40	0.00	0.00 - NA
rs17820943	*MAFB*	0.93	1.01	0.75 - 1.36	0.72	0.94	0.68 - 1.30
rs13041247	*MAFB*	0.80	1.04	0.77 - 1.40	0.99	1.00	0.72 - 1.38
rs11696257	*MAFB*	0.87	1.02	0.76 - 1.38	0.76	0.95	0.69 - 1.31

[a]Loci that reached nominal significance

Table S4: Case-control analyses for Nigeria

Part A: Case-control analyses for NSCL/P and NSCPO for Nigeria							
	Probable	NSCL/P			NSCPO		
SNP	gene/loci	*p*	OR	95% CI	*p*	OR	95% CI
rs1801131	*MTHFR*	0.42	1.17	0.80 - 1.70	0.28	0.62	0.26 - 1.47
rs1801133	*MTHFR*	0.07	1.53	0.96 - 2.45	0.73	0.83	0.29 - 2.37
rs766325	*PAX7*	0.56	0.90	0.64 - 1.27	0.07	0.50	0.27 - 1.07
rs742071	*PAX7*	0.05	1.30	1.00 - 1.67	0.93	1.02	0.65 - 1.62
rs560426	*ABCA4*	0.77	0.96	0.75 - 1.24	0.94	0.98	0.63 - 1.55
rs481931	*ABCA4*	0.12	1.40	0.92 - 2.12	0.89	0.94	0.39 - 2.25
rs4147811	*ABCA4*	**7.48E-03**[a]	**1.72**	**1.15 - 2.56**	0.44	1.34	0.64 - 2.80
rs138751793	*ARHGAP29*	0.12	1.69	0.86 - 3.32	0.57	1.43	0.41 - 4.93
rs6677101	*SLC25A24*	0.66	0.94	0.72 - 1.23	0.08	0.63	0.38 - 1.06
rs861020	*IRF6*	0.90	1.02	0.70 - 1.49	0.78	0.91	0.47 - 1.77
rs34743335	*IRF6*	0.28	2.16	0.51 - 9.08	0.06	9.33	0.58 - 150.60
rs642961	*IRF6*	0.60	0.89	0.57 - 1.38	0.67	0.85	0.39 - 1.82
rs7590268	*THADA*	0.89	0.98	0.71 - 1.35	0.66	0.87	0.48 - 1.60
rs4332945	*DYSF*	0.51	1.13	0.79 - 1.60	0.55	1.21	0.66 - 2.22
rs2303596	*DYSF*	0.14	1.28	0.92 - 1.77	0.21	0.63	0.31 - 1.30
rs227782	*DYSF*	0.79	0.97	0.74 - 1.25	0.29	0.77	0.48 - 1.24
rs115200552	*MSX1*	0.23	0.59	0.24 - 1.41	0.15	1.93	0.78 - 4.79
rs12532	*MSX1*	0.49	0.91	0.71 - 1.18	0.15	0.72	0.45 - 1.13
rs2674394	Gene Desert	0.81	1.04	0.74 - 1.47	0.97	1.01	0.54 - 1.89
rs651333	*TULP4*	0.51	0.91	0.68 - 1.21	0.12	1.45	0.91 - 2.32
rs6558002	*EPHX2*	0.92	0.98	0.71 - 1.36	0.80	1.08	0.62 - 1.86
rs987525	8q24	0.74	0.96	0.73 - 1.25	0.31	0.78	0.49 - 1.26
rs894673	*FOXE1*	0.49	1.10	0.84 - 1.46	0.73	1.09	0.67 - 1.79
rs3758249	*FOXE1*	0.40	1.13	0.85 - 1.18	0.74	1.09	0.66 - 1.78
rs7078160	*VAX1*	0.63	1.07	0.80 - 1.44	1.00	1.00	0.59 - 1.70
rs4752028	*VAX1*	0.68	1.06	0.82 - 1.36	0.98	0.99	0.63 - 1.57
rs10785430	*ADAMTS20*	0.77	1.04	0.79 - 1.36	0.67	0.90	0.55 - 1.47
rs9574565	*SPRY2*	0.57	0.92	0.70 - 1.21	0.77	0.93	0.58 - 1.49
rs8001641	*SPRY2*	0.80	0.94	0.58 - 1.52	0.75	1.13	0.52 - 2.45
rs17563	*BMP4*	0.29	1.21	0.85 - 1.72	0.38	1.31	0.72 - 2.37
rs1258763	*GREM1*	0.46	1.10	0.85 - 1.42	0.84	0.95	0.61 - 1.50
rs8049367	*ADCY9*	0.09	1.26	0.97 - 1.64	0.11	0.66	0.39 - 1.11
rs16260	*CDH1*	0.78	1.06	0.71 - 1.58	0.53	0.78	0.37 - 1.68
rs11642413	*CDH1*	0.62	0.93	0.69 - 1.25	0.33	0.75	0.42 - 1.34
rs1546124	*CRISPLD2*	0.62	0.93	0.68 - 1.26	0.51	0.83	0.48 - 1.44
rs4783099	*CRISPLD2*	0.98	1.00	0.77 - 1.31	**0.04**[a]	**0.58**	**0.34 - 0.98**
rs8069536	*NTN1*	0.06	1.28	0.99 - 1.66	0.62	1.12	0.71 - 1.78

rs8081823	*NTN1*	0.30	0.85	0.63 - 1.15	0.67	0.89	0.53 - 1.51
rs17760296	*NOG1*	0.29	1.45	0.73 - 2.88	0.11	2.21	0.81 - 6.02
rs227731	*NOG1*	0.83	0.97	0.70 - 1.34	0.56	1.17	0.68 - 2.02
rs7224837	*AXIN2*	0.71	0.92	0.60 - 1.42	0.67	1.16	0.58 - 2.35
rs3923086	*AXIN2*	0.35	0.00	0.00 - NA	0.56	0.00	0.00 - NA
rs17820943	*MAFB*	0.81	1.04	0.76 - 1.42	0.21	1.38	0.83 - 2.30
rs13041247	*MAFB*	0.81	1.04	0.76 - 1.42	0.22	1.38	0.83 - 2.30
rs11696257	*MAFB*	0.78	1.05	0.77 - 1.43	0.21	1.38	0.83 - 2.31
Part B: Case-control analyses for NSCL/P subphenotypes for Nigeria							
SNP	Probable gene/loci	NSCL			NSCLP		
		p	OR	95% CI	*p*	OR	95% CI
rs1801131	*MTHFR*	0.61	1.14	0.68 - 1.91	0.36	1.27	0.76 - 2.10
rs1801133	*MTHFR*	0.17	1.56	0.82 - 2.95	0.15	1.59	0.84 - 3.00
rs766325	*PAX7*	0.87	0.96	0.61 - 1.52	0.38	0.80	0.49 - 1.31
rs742071	*PAX7*	**0.02**[a]	**1.48**	**1.05 - 2.08**	0.27	1.22	0.86 - 1.74
rs560426	*ABCA4*	0.98	1.01	0.71 - 1.41	0.63	1.09	0.77 - 1.56
rs481931	*ABCA4*	0.66	0.86	0.43 - 1.71	**2.87E-03**[a]	**2.10**	**1.28 - 3.46**
rs4147811	*ABCA4*	**0.02**[a]	**1.88**	**1.11 - 3.18**	0.05	1.72	1.00 - 2.95
rs138751793	*ARHGAP29*	**0.04**[a]	**2.30**	**1.04 - 5.09**	0.80	1.15	0.39 - 3.39
rs6677101	*SLC25A24*	0.88	1.03	0.72 - 1.47	0.35	0.83	0.57 - 1.22
rs861020	*IRF6*	0.59	0.87	0.52 - 1.46	0.87	1.04	0.63 - 1.74
rs34743335	*IRF6*	0.38	2.79	0.25 - 30.91	0.16	3.20	0.58 - 17.63
rs642961	*IRF6*	0.20	0.66	0.34 - 1.26	0.97	1.01	0.57 - 1.80
rs7590268	*THADA*	0.94	1.02	0.67 - 1.55	0.87	0.96	0.61 - 1.52
rs4332945	*DYSF*	0.63	1.13	0.70 - 1.81	0.62	1.13	0.70 - 1.84
rs2303596	*DYSF*	0.89	0.97	0.60 - 1.57	0.05	1.51	0.99 - 2.31
rs227782	*DYSF*	0.88	1.03	0.73 - 1.45	0.76	0.95	0.66 - 2.56
rs115200552	*MSX1*	0.16	0.37	0.09 - 1.56	0.68	0.80	0.28 - 2.30
rs12532	*MSX1*	0.25	0.82	0.58 - 1.15	0.82	0.96	0.67 - 1.37
rs2674394	Gene Desert	0.57	1.14	0.73 - 1.80	0.97	0.99	0.62 - 1.59
rs651333	*TULP4*	0.40	0.85	0.58 - 1.25	0.93	0.98	0.67 - 1.45
rs6558002	*EPHX2*	0.53	0.87	0.55 - 1.36	0.86	1.04	0.67 - 1.63
rs987525	8q24	0.44	0.87	0.61 - 1.24	0.81	0.96	0.66 - 1.38
rs894673	*FOXE1*	0.15	1.30	0.91 - 1.87	0.78	0.94	0.63 - 1.41
rs3758249	*FOXE1*	0.14	1.31	0.91 - 1.87	0.89	0.97	0.65 - 1.44
rs7078160	*VAX1*	0.17	1.30	0.89 - 1.89	0.37	0.82	0.53 - 1.27
rs4752028	*VAX1*	0.42	1.15	0.82 - 1.62	0.96	1.01	0.71 - 1.44
rs10785430	*ADAMTS20*	0.53	0.89	0.61 - 1.29	0.27	1.23	0.85 - 1.77
rs9574565	*SPRY2*	0.90	0.98	0.68 - 1.40	0.34	0.83	0.57 - 1.22
rs8001641	*SPRY2*	0.71	0.88	0.47 - 1.67	0.85	0.94	0.48 - 1.82
rs17563	*BMP4*	0.17	1.38	0.87 - 2.18	0.64	1.12	0.69 - 1.82

rs1258763	*GREM1*	0.64	1.09	0.77 - 1.53	0.31	1.20	0.84 - 1.71
rs8049367	*ADCY9*	0.09	1.36	0.96 - 1.93	0.45	1.15	0.80 - 1.66
rs16260	*CDH1*	0.75	1.09	0.65 - 1.82	0.85	0.94	0.53 - 1.68
rs11642413	*CDH1*	0.31	0.81	0.53 - 1.23	0.50	1.15	0.77 - 1.71
rs1546124	*CRISPLD2*	0.49	0.87	0.57 - 1.31	0.82	0.95	0.62 - 1.45
rs4783099	*CRISPLD2*	0.73	1.07	0.75 - 1.52	0.83	0.96	0.66 - 1.39
rs8069536	*NTN1*	0.19	1.26	0.89 - 1.78	0.12	1.33	0.93 - 1.92
rs8081823	*NTN1*	0.85	0.96	0.65 - 1.42	0.10	0.70	0.45 - 1.08
rs17760296	*NOG1*	0.19	1.71	0.76 - 3.86	0.80	1.15	0.39 - 3.40
rs227731	*NOG1*	0.31	0.79	0.50 - 1.25	0.44	1.18	0.77 - 1.82
rs7224837	*AXIN2*	0.97	1.01	0.58 - 1.78	0.61	0.85	0.46 - 1.57
rs3923086	*AXIN2*	0.46	0.00	0.00 - NA	0.48	0.00	0.00 - NA
rs17820943	*MAFB*	0.91	1.02	0.67 - 1.56	0.62	1.11	0.73 - 1.70
rs13041247	*MAFB*	0.94	1.02	0.67 - 1.54	0.61	1.12	0.73 - 1.70
rs11696257	*MAFB*	0.91	1.02	0.67 - 1.56	0.59	1.12	0.74 - 1.71

[a]Loci that reached nominal significance

Table S5: Other rare and/or potentially aetiologic variants observed in seven sequenced genes

HGVS	HGVp	α	¶	b	Polyphen-2	SIFT	§	¥	Reference
ARHGAP29									
c.560-199T>C	N/A	1	NSCLP	N/A	N/A	N/A	β	N/A	dbSNP
c.1144-18T>C	N/A	2	1 NSCLP and I NSCL	N/A	N/A	N/A	β,μ	N/A	dbSNP
c.2738C>T	p.Ser913Leu	4	2 NSCLP, 1 NSCL and 1 CPO	4 d	Benign	Deleterious	N/A	N/A	dbSNP
c.2957T>C	p.Ile986Thr	1	NSCLP	N/A	Benign	Tolerated	N/A	N/A	dbSNP
c.2962G>T	p.Asp988Tyr	2	NSCLP	1 d, 1 g	Probably Damaging	Deleterious	N/A	N/A	dbSNP
c.3023G>A	p.Arg1008Lys	2	1 NSCLP and 1 CPO	N/A	Benign	Tolerated	N/A	N/A	dbSNP
VAX1									
c.390G>A	p.Arg130Arg	1	NSCPO	N/A	Benign	Tolerated	ε	N/A	Novel
c.429+37G>C	N/A	1	NSCLP	N/A	N/A	N/A	β	N/A	1000Genome
c.429+50C>A	N/A	4	1 NSCLP, 1 NSCL and 2 CPO	N/A	N/A	N/A	μ	λ	1000Genome
c.693C>A	p.Ala231Ala	4	1 NSCLP and 3 NSCPO	N/A	Benign	Tolerated	γ, ε	λ	Novel
c.754G>T	p.Gly252Cys	1	NSCL	e	Probably	Deleterious	N/A	N/A	Novel

					Damaging				
PAX7									
c.703G>A	p.Ala235Thr	2	NSCLP	1 c, 1 d	Probably Damaging	Deleterious	N/A	N/A	dbSNP and ExAC
c.1223C>T	p.Pro408Leu	1	CPO	d	Probably Damaging	Deleterious	N/A	N/A	dbSNP and ExAC
MSX1									
c.95C>T	p.Ala32Val	4	2 NSCL and 2 CPO	N/A	Benign	Tolerated	η,ε	N/A	ExAc
c.218C>T	p.Pro73Leu	3	NSCL	2 d, 1 f	Possibly Damaging	Deleterious	N/A	N/A	dbSNP
c.522G>A	p.Lys174Lys	1	NSCL	N/A	Benign	Tolerated	N/A	N/A	Novel
BMP4									
c.860G>A	p.Arg287His	1	NSCL	N/A	Benign	Tolerated	N/A	N/A	dbSNP
c.371-164G>A	N/A	2	1 NSCLP and 1 NSCL	N/A	N/A	N/A	β, μ	N/A	Novel
c.280G>A	p.Glu94Lys	1	NSCL	N/A	Benign	Tolerated	N/A	N/A	Novel
c.228T>A	p.Ser76Arg	3	2 NSCLP and 1 NSCL	1 d, 2 f	Possibly Damaging	Damaging to two isoforms	N/A	N/A	dbSNP
FOXE1									

c.107C>T	p.Thr36Met	1	NSCLP	d	Possibly Damaging	Deleterious	N/A	N/A	ExAc
c.569C>G	p.Pro190Arg	6	3 NSCLP, 2 NSCL and 1 CPO	1 c, 2 d, 1 e, 2 f	Possibly Damaging	Deleterious	N/A	N/A	dbSNP
MAFB									
c.-1G>A	N/A	2	1 NSCL and 1 CPO	N/A	N/A	N/A	δ	N/A	dbSNP
c.-75C>A	N/A	1	NSCLP	N/A	N/A	N/A	N/A	λ	dbSNP
c.603G>C	p.Ala201Ala	1	NSCL	N/A	Benign	Tolerated	N/A	N/A	Novel

α: Total number of cases with variant, **¶**: subphenotype of probands in which the variant was observed, **b**: segregation analyses, **c**: variant was observed in clinically unaffected father, **d**: variant was observed in clinically unaffected mother, **e**: variant was absent in the only paternal sample available, **f**: variant was absent in the only maternal sample available, **g**: no parental sample was available, **§**: Human Splice Finder, **¥:** RegulomeDB, **β:** alteration of Exonic Splicing Silencer (ESS) Site, **μ:** creation of new Exonic Splicing Enhancer (ESE) site, **δ:** Alteration of the wildtype (WT) donor site, **ε:** alteration of an ESE site, **γ**: creation of new Acceptor site with branch points, **η:** creation of new Donor site, **λ:** 2b - Likely to affect binding of *EZH2*, **N/A:** Not Applicable, **NSCLP**: nonsyndromic cleft lip and palate, **NSCL**: nonsyndromic cleft lip only, **CPO**: cleft palate only. All analyses were based on genome assembly number GRCh37/hg19, 2009 (http://genome.ucsc.edu).

Department of Biochemistry and Biotechnology,

Kwame Nkrumah University of Science and Technology,

PMB, University Post Office, Kumasi, Ghana.

6th June 2016

Editor-in-Chief,

Journal of Dental Research

Dear Sir,

Re: "Association studies and direct DNA sequencing implicate some known genetic susceptibility loci in the etiology of nonsyndromic orofacial clefts in sub-Saharan African populations": JDR-16-0113.R2

We are submitting a revised version of the above manuscript to your journal for publication. This revision is in response to the reviews' comments. As requested by the reviewers, we have now added national prevalence data for Nigeria (though none exists for Ghana and Ethiopia) in the introduction. In the discussion, we have also hypothesized the possible existence of protective genetic variants in the African genome that may reduce OFC susceptibility.

Responses to reviewers' comments are shown in the next page. We thank the reviewers for their comments and the Editor for the opportunity to make these revisions and for us to be able to send our responses.

Yours faithfully

Lord Jephthah Joojo Gowans

Response to Reviewers' comment

Thank you for making the revisions to your very interesting manuscript. The edits to the tables and results are definitely an improvement. I just have one more request. Would you please add in the incidence of OFC types in Africans? Right now you have only stated the global incidence but this is less relevant for your study. The reader needs a bit more context. I for one, would like to know if it is still true to say that the incidence of NSCLP is lower in Africans. This information plus appropriate citation should be added to the introduction. In the discussion, it would be important to reference the incidence especially if it is much different than for Europeans. How would you reconcile the idea that the incidence is lower but the same genetic variants are involved? Does this mean there are protective variants somewhere in the genome? It would be worth mentioning this.

Response: we have added "Though there is no national prevalence data for Ghana and Ethiopia, a prevalence estimate of 0.5 per 1000 has been observed for Nigeria (Butali et al. 2014a)" to the introduction.

We have dedicated the last paragraph of the discussion to elucidating the implications of the lower incidence in Africans, though Africans may share similar or same genetic susceptibility variants with Asians and Europeans.